



UNIVERSIDADE FEDERAL DE OURO PRETO
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS
DEPARTAMENTO DE CIÊNCIAS BIOLÓGICAS
LABORATÓRIO DE BIOLOGIA E TECNOLOGIA DE
MICROORGANISMOS



UBIRATAN DA SILVA BATISTA

**Triagem otimizada de epítomos consensuais do antígeno estrutural E2 de
CHIKV e EEEV: Uma abordagem *in silico***

OURO PRETO – MG

2022

UBIRATAN DA SILVA BATISTA

**Triagem otimizada de epítomos consensuais do antígeno estrutural E2 de
CHIKV e EEEV: Uma abordagem *in silico***

Trabalho de conclusão de curso apresentado a
Universidade Federal de Ouro Preto (UFOP), como
parte dos requisitos para a obtenção do título de
Bacharel em Ciências Biológicas.

Área de concentração: Biotecnologia

Orientador: Prof. Dr. Breno de Mello Silva

Co-orientador: Msc. Ricardo Lemes Gonçalves

OURO PRETO – MG

2022

SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

B333t Batista, Ubiratan Da Silva.
Triagem otimizada de epítomos consensuais do antígeno estrutural E2 de CHIKV e EEEV [manuscrito]: uma abordagem in silico. / Ubiratan Da Silva Batista. - 2022.
62 f.: il.: color., tab..

Orientador: Prof. Dr. Breno de Mello Silva.
Coorientador: Me. Ricardo Lemes Gonçalves.
Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Instituto de Ciências Exatas e Biológicas. Graduação em Ciências Biológicas .

1. Vacinas. 2. Múltiplos-epítomos. 3. Epítomos lineares. 4. Alphavirus.
5. Antígenos. I. Gonçalves, Ricardo Lemes. II. Silva, Breno de Mello. III.
Universidade Federal de Ouro Preto. IV. Título.

CDU 615.371

Bibliotecário(a) Responsável: Luciana De Oliveira - SIAPE: 1.937.800



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DE OURO PRETO
REITORIA
INSTITUTO DE CIÊNCIAS EXATAS E BIOLÓGICAS
DEPARTAMENTO DE CIÊNCIAS BIOLÓGICAS



FOLHA DE APROVAÇÃO

Ubiratan da Silva Batista

Triagem otimizada de epítomos consensuais do antígeno estrutural E2 de CHIKV e EEEV: Uma abordagem *in silico*

Monografia apresentada ao Curso de Ciências Biológicas da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de bacharel Ciências Biológicas.

Aprovada em 23 de junho de 2022.

Membros da banca

Dr. B.M.S. - Orientador - Universidade Federal de Ouro Preto

Msc. R.L.G. - Co-orientador - Universidade Federal de Ouro Preto

Dra. L.M.C. - Universidade de São Paulo

Dr. R.A.M.A. - Universidade do Extremo Sul Catarinense

BMS, orientador do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 23/06/2022.



Documento assinado eletronicamente por **Breno de Mello Silva, PROFESSOR DE MAGISTERIO SUPERIOR**, em 04/10/2022, às 16:54, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0407744** e o código CRC **8A3A6B71**.

Referência: Caso responda este documento, indicar expressamente o Processo nº 23109.008169/2022-19

SEI nº 0407744

R. Diogo de Vasconcelos, 122, - Bairro Pilar Ouro Preto/MG, CEP 35400-000
Telefone: 3135591672 - www.ufop.br

Criado por [03163003656](#), versão 2 por [03163003656](#) em 04/10/2022 16:54:16.

AGRADECIMENTOS

À Deus por todas as bênçãos e dádivas a mim concedidas.

Ao povo brasileiro pela manutenção da escola pública, gratuita e de qualidade.

Às vítimas da COVID-19.

À minha família por todo amor, carinho e cumplicidade: paiinho, Adolfo Mendes Batista; mainha, Maria Isaura da Silva Batista; e irmãos Jacira, Felipe, Pedro, Tainan e Ubirajara.

Às minhas avós Hilda e Nem (in memoriam) e avô Fidelis, pelas orações e incentivo.

Aos tios(as), primos(as) e amigos(as) que sempre se fizeram presentes no meu caminho e em minha torcida.

Ao professor orientador Dr. Breno de Mello Silva pela orientação e oportunidade.

Ao coorientador Ricardo Lemes Gonçalves pela parceira, pelos ensinamentos, pelos cafés filosóficos e pela brilhante coorientação.

À Ana Clara Gomes, minha parceira de Iniciação Científica, pelas contribuições no trabalho e por me convidar a embarcar nas loucuras da bioinformática.

Aos alunos de pós-graduação do meu grupo de pesquisa: Bruna, Camila, Cyntia, Júlia, Orlando e Túlio, pelas discussões e pelas contribuições no trabalho.

À república Mata Virgem, família para vida toda, pela amizade e parceria em todos os anos da graduação. A jornada foi mais prazerosa e especial com vocês!

À Daíze Maria de Jesus, pela amizade, incentivo e escuta, mesmo de longe.

À minha biogalera: Alzira, Raquel, Ana, João, Karol, Laysi e Diego pela amizade, incentivo, companheirismo, rocks e parceria nos mais variados momentos dessa jornada.

À Ouro Preto, cidade que me acolheu e foi minha segunda casa.

À UFOP pela infraestrutura na realização do trabalho e a todos os mestres da casa, pelos ensinamentos adquiridos.

“O sonho é que leva a gente para a frente. Se a gente for seguir a razão, fica aquietado, acomodado.”

Ariano Suassuna

RESUMO

Nos últimos anos, a vacinologia reversa e a imunoinformática revolucionaram a construção de protótipos vacinais. Apesar dos avanços nas etapas de seleção de alvos imunogênicos, doenças virais importantes, como as causadas por Alphavirus, permanecem sem tratamentos ou vacinas licenciadas nas principais agências de saúde do mundo. No presente estudo, nós identificamos através de abordagens *in silico* epítomos lineares e conservados das proteínas estruturais E2 dos Alphavirus *Chikungunya virus* (CHIKV) e *Eastern equine encephalitis virus* (EEEV). Para identificação dos epítomos foram utilizadas as predições das ferramentas: Bepipred 2.0 e Predicted Antigenic Peptides/IMED para células B, NetCTL para células T-citotóxicas e IEDB – Binding Predictions para células T-auxiliares. Para predição de conservação foi usada a ferramenta IEDB - Epitope Conservancy Analysis. Foram identificados 294 peptídeos preditadamente antigênicos nas proteínas E2 destes vírus, sendo 71,20% para células T-auxiliares, 19,92% para células B e 8,87% para células T-citotóxicas. Com 60% de identidade mínima entre as sequências, a análise de conservação identificou 63 (21.50%) destes peptídeos como sendo conservados. Desconsiderando a redundância das sequências, 40 destes epítomos foram selecionados. Os epítomos conservados foram analisados na estrutura tridimensional e estão distribuídos entre os 3 domínios da proteína. A identificação dos alvos conservados entre Alphavirus artritogênicos e encefalitogênicos constitui importante via para o desenho otimizado de antígenos para plataformas vacinais. Além disso, a seleção otimizada destes epítomos pelas ferramentas POA1 e POA2, mostram que estas ferramentas podem ser usadas para identificação de peptídeos em múltiplas abordagens visando proteínas virais para o desenvolvimento de vacinas e testes diagnósticos.

Palavras-chave: vacinas, imunoinformática, múltiplos-epítomos, epítomos lineares, Alphavirus

ABSTRACT

In recent years, reverse vaccinology and immunoinformatics have revolutionized the construction of vaccine prototypes. Despite advances in immunogenic targeting steps, relevant viral diseases, such as those caused by Alphavirus, remain without treatments or vaccines licensed by the world's leading health agencies. In the present study, we identified linear and conserved epitopes of the E2 structural proteins of Alphavirus Chikungunya virus (CHIKV) and Eastern equine encephalitis virus (EEEV) using *in silico* approaches. The screening of epitopes was carried out by the use of prediction tools: Bepipred 2.0 and Predicted Antigenic Peptides/IMED for B cells, NetCTL for T-cytotoxic cells and IEDB – Binding Predictions for T-helper cells. The tool IEDB - Epitope Conservancy Analysis was used for conservancy analysis. A total of 294 predicted antigenic peptides were identified in the E2 proteins of these viruses, 71.20% for T-helper cells, 19.92% for B cells and 8.87% for T-cytotoxic cells. With 60% minimal sequence identity, conservation analysis identified 63 (21.50%) of these peptides as being conserved. Disregarding sequence redundancy, 40 of these epitopes were considered. The conserved epitopes were analyzed in the three-dimensional structure and are distributed among the 3 domains of the protein. The identification of conserved targets between arthritogenic and encephalitogenic Alphaviruses constitutes an important route for the optimized design of antigens for vaccine platforms. Furthermore, the optimized selection of these epitopes by the POA1 and POA2 tools show that these tools can be used for peptide identification in multiple approaches targeting viral proteins for vaccine development and diagnostic tests.

Keywords: vaccines, immunoinformatics, multi-epitopes, linear epitopes, Alphavirus

LISTA DE FIGURAS

Figura 1: Organização geral do genoma dos Alphavirus.....	pág. 13
Figura 2: Relações filogenéticas de <i>Alphavirus</i> de interesse médico.....	pág. 14
Figura 3: Imagem de cristal de crio-microscopia eletrônica da partícula de CHIKV.....	pág. 16
Figura 4: Proteínas do Envelope em vista tridimensional.....	pág. 17
Figura 5: Fluxograma das etapas da Triagem Otimizada de Antígenos presentes no Pipeline de Otimização de Antígenos (POA).....	pág. 27
Figura 6: Comparação entre as etapas da seleção de epítomos realizadas manualmente e com o uso da triagem otimizada do POA.....	pág. 30
Figura 7: Resultados da validação da estrutura tridimensional da proteína E2 de CHIKV.....	pág. 36
Figura 8: Resultados da validação da estrutura tridimensional da proteína E2 de EEEV.....	pág. 37
Figura 9: Visualização 3D da distribuição dos epítomos preditos na proteína E2 de EEEV.....	pág. 38

LISTA DE TABELAS

- Tabela 1:** Quantidades de epítomos da proteína E2 dos *Alphavirus* CHIKV e EEEV preditos para células B e células T.....pág. 31
- Tabela 2:** Quantidade de epítomos conservados entre as proteínas E2 dos *Alphavirus* CHIKV e EEEV.....pág. 32
- Tabela 3:** Epítomos conservados das proteínas E2 de CHIKV e EEEV, triados com base nas ferramentas de predição de epítopo, de análise de conservação e de predição de porções transmembranares.....pág. 33

LISTA DE ABREVIATURAS

NCBI - National Center for Biotechnology Information

CHIKV - *Chikungunya virus*

EEEV - *Eastern equine encephalitis virus*

MAYV - *Mayaro virus*

ONNV - *Onyong-nyong virus*

RRV - *Ross River virus*

SFV - *Semliki Forest virus*

SINV - *Sindbis virus*

VEEV - *Venezuelan equine encephalitis virus*

WEEV - *Western equine encephalitis virus*

SUMÁRIO

1. INTRODUÇÃO	10
2. REVISÃO BIBLIOGRÁFICA.....	13
2.1 Arboviroses.....	13
2.2 Alphavirus: Estrutura e Características	14
2.3 Partícula Viral e Proteínas estruturais	15
2.4 Peptídeos Imunogênicos	17
2.5 Nova era das vacinas e vacinologia reversa	18
2.6 Abordagens <i>in silico</i> e Imunoinformática	19
2.7 Predições de epítomos	20
3. OBJETIVOS.....	21
3.1 Objetivo geral:	21
3.2 Objetivos específicos:.....	21
4. MATERIAIS E MÉTODOS	22
4.1 Aquisição de Sequências	22
4.2 Predição de epítomos.....	22
4.3 Triagem dos resultados das predições de epítomos.....	24
4.4 Análise de Conservação.....	25
4.5 Triagem dos resultados da análise de conservação e análise de propensão a regiões transmembranares.....	25
4.6 Análise de Toxicidade e Alergenicidade	26
4.7 Caracterização Tridimensional dos Alvos.....	27
4.8 Cálculo de acessibilidade ao Solvente.....	28
5. RESULTADOS E DISCUSSÕES	30
6. CONSIDERAÇÕES FINAIS	40
7. REFERÊNCIAS BIBLIOGRÁFICAS	41

8. ANEXOS.....	54
8.1 Sequências das proteínas E2 dos vírus CHIKV e EEEV.....	54
8.2 Arquivo FASTA criado a partir do resultado da predição da ferramenta PAP/IMED..	55
8.3 Arquivo Fasta retornado pela ferramenta POA1 para epítomos da proteína E2 de CHIKV	56
8.4 Alinhamento entre as sequencias das proteínas E2 dos Alphavirus CHIKV e EEEV	57
8.5 Alinhamento múltiplo entre as sequencias das proteínas E2 dos Alphavirus CHIKV e EEEV e os epítomos selecionados pela Triagem Otimizada de Antígenos.....	58
8.6 Epítomos de Linfócitos T auxiliares preditos para mais de um alelo de HLA pela ferramenta MHCII – Binding Predictions.....	59
8.7 Visualização dos epítomos selecionados pela triagem otimizada do POA em estrutura tridimensional da proteína E2 de CHIKV	61
8.8 Área exposta ao solvente dos Epítomos de células B selecionados pela Triagem Otimizada de Antígenos	62

1. INTRODUÇÃO

O gênero *Alphavirus*, único pertencente à família *Togaviridae*, corresponde a 31 espécies de vírus, muitos deles arbovírus distribuídos mundialmente (CHEN et al., 2018; MENDES; KUHN, 2018). Os *Alphavirus* são vírus envelopados, de RNA de fita simples sense-positivo capazes de infectar uma gama de hospedeiros, desde aves à mamíferos (CONWAY; COLPITTS; FIKRIG, 2014; MENDES; KUHN, 2018; RANGEL et al., 2021).

Embora de interesse médico, alguns destes vírus não são considerados grandes problemas de saúde pública por apresentarem baixa mortalidade, área de influência restrita ou por manterem sua transmissão e ciclo de vida somente no meio natural, entre primatas não-humanos e outros organismos (DIAGNE et al., 2020; SUHRBIER; JAFFAR-BANDJEE; GASQUE, 2012; ZACKS; PAESSLER, 2010). No entanto, segundo a Organização Mundial de Saúde (OMS, 2022), as mudanças climáticas e as alterações no ambiente natural têm promovido um quadro propício para que doenças virais emergentes e re-emergentes causem epidemias globais (LUNDSTROM, 2014; OLIVIA et al., 2015).

Dentre os *Alphavirus* de interesse médico, destaca-se o vírus *Chikungunya Virus* (CHIKV) que promoveu surtos mais recentes, em 2016/17 (WHO, 2020b). A febre do Chikungunya apresenta sintomas muito parecidos aos da febre da Dengue, embora a mortalidade seja de 0,1% para esta primeira e de 0,5%-3,5% para a segunda (FURUYA-KANAMORI et al., 2016). Tais febres apresentam um grande risco de mortalidade para crianças menores de 1 ano e adultos com mais de 60 anos (LANGSJOEN et al., 2018). Segundo a Organização Mundial de Saúde, os *Alphavirus* CHIKV e *Mayaro virus* (MAYV) causam milhares de infecções todos os anos e os países em desenvolvimento na Ásia e nas Américas são os mais afetados. Em 2017, por exemplo, a doença causada pelo CHIKV acometeu 548 pessoas na Europa, >100.000 pessoas na Ásia e 185 mil nas américas e no Caribe (WHO, 2020b).

Um outro *Alphavirus* de interesse médico, o vírus *Eastern equine encephalitis virus* (EEEV), também tem sido reportado com importância médica e veterinária. Com casos mais frequentemente relatados nas Américas, estes vírus podem causar infecções graves em equinos, em algumas espécies de aves e em humanos (CARRERA et al., 2013; ZACKS; PAESSLER, 2010). Embora, na maioria dos casos, seja associado à uma doença autolimitada, o EEEV pode

causar encefalite grave com mortalidade entre 30 e 75% em humanos - a depender da idade, e de até 90% em cavalos. A infecção pode ainda ocasionar defeitos neurológicos duradouros em indivíduos sobreviventes, podendo chegar a 90% os níveis de dano (STROMBERG et al., 2020).

Assim como para outras doenças tropicais negligenciadas, apesar da importância médica e epidemiológica, até os dias atuais não existem tratamentos ou vacinas licenciadas para o combate a doenças causadas por *Alphavirus* nas principais agências de saúde do mundo (GAO; SONG; ZHANG, 2019; SHARMA; KNOLLMANN-RITSCHER, 2019; STROMBERG et al., 2020). Embora existam candidatos vacinais em diferentes etapas clínicas de teste (WEISS et al., 2020), investimentos no desenvolvimento de vacinas contra arboviroses tem sido, de maneira geral, associados à surtos e não como forma estratégica de proteção da população contra doenças re-emergentes (STROMBERG et al., 2020).

Nas últimas décadas, o surgimento de novas tecnologias de sequenciamento genético e os avanços em pesquisa e produção de vacinas revolucionaram as técnicas de construção de protótipos vacinais, cada vez mais rápidas e menos dispendiosas (BAE et al., 2009; URA et al., 2021). Neste cenário, surgiram novas perspectivas para o uso da tecnologia na produção de vacinas, como a imunoinformática e a vacinologia reversa. Estes avanços permitiram a identificação de antígenos a partir das informações genômicas dos patógenos e redesenharam a descobertas de epítomos antigênicos capazes de provocar resposta imune celular e/ou humoral (SETTE; RAPPUOLI, 2010).

Uma abordagem crescente no contexto de novas plataformas vacinais tem sido as vacinas de múltiplos-epítomos. Esta tecnologia tem se mostrado uma estratégia promissora na construção de candidatos vacinais baseados em peptídeos contra tumores e doenças causadas por infecções virais (ZHANG, 2017). Utilizando diferentes ferramentas de biologia computacional e imunoinformática, as plataformas de múltiplos-epítomos consistem na seleção racional de epítomos e na construção, a partir destes peptídeos candidatos, de modelos quiméricos de proteínas antigênicas capazes de elicitar simultaneamente fortes respostas imunes celulares e humorais (ALMOFTI; ABD-ELRAHMAN; ELTILIB, 2021; JIANG et al., 2017; SHEY et al., 2019).

Contudo, o sucesso das plataformas de múltiplos-epítomos está intimamente ligado a seleção apropriada dos epítomos para a construção dos modelos quiméricos. Embora os métodos de predição disponíveis possam identificar epítomos potenciais em qualquer proteína, nem todos eles serão relevantes para o desenvolvimento de vacinas (SANCHEZ-TRINCADO; GOMEZ-PEROSANZ; RECHE, 2017). Diante disso, o presente trabalho visa encontrar em abordagem *in silico* epítomos alvos em antígenos de *Alphavirus* de interesse médico que sejam competentes para a construção de um candidato vacinal. As abordagens de imunoinformática adotadas estão em consonância às técnicas utilizadas na literatura e aos trabalhos prévios publicados pelo grupo (GONÇALVES et al., 2020), bem como utilizam de ferramentas computacionais desenvolvidas no âmbito deste trabalho, atualmente validadas para a triagem, seleção e coleta de epítomos (POA – LBTM. https://github.com/UbiratanBatista/POA_Project).

As características selecionadas nos epítomos foram a imunogenicidade, a linearidade, a conservação entre as espécies e a topologia de membrana. Tais características são preconizadas para a construção de uma vacina eficaz e barata, uma vez capaz de ser produzida em sistemas celulares mais simples, como o bacteriano.

2. REVISÃO BIBLIOGRÁFICA

2.1 Arboviroses

Arboviroses são doenças causadas por vírus que são transmitidos por artrópodes (insetos, aracnídeos, etc.), também chamados arbovírus. Esta classificação é bem diversa e engloba, por exemplo, os vírus da Dengue (DENV), Zika (ZIKAV), Chikungunya (CHIKV), Febre Amarela (YFV), além de meningite e encefalites virais (FIGUEIREDO; PAIVA; MORATO, 2017).

Geralmente, estes são vírus que possuem um ciclo de vida silvestre, onde os vetores artrópodes (ex. *Aedes aegypti*, *Aedes albopictus*, etc.) se alimentam de espécies da fauna silvestre em regiões de mata e/ou savana e o ciclo de vida deste vírus ocorre sendo transmitido de um hospedeiro ao outro, sem que os hospedeiros ou os vetores apresentem evidência clínica de infecção pelo vírus (GOULD; HIGGS, 2009). Porém, pelo menos 150 espécies de arbovírus são conhecidos por infectarem humanos e/ou animais domésticos, podendo causar quadros sintomáticos variados (CAICEDO et al., 2021; GAO; SONG; ZHANG, 2019; SHARMA; KNOLLMANN-RITSCHER, 2019; WEAVER; REISEN, 2010).

Em geral, por serem vírus de RNA, os arbovírus possuem grande plasticidade genética que os permitem executar ciclos de replicações alternados entre hospedeiros vertebrados e invertebrados. Alguns destes vírus, como os *Alphavirus* e os *Flavivirus*, possuem grande habilidade de dispersão utilizando-se de vetores hoje considerados cosmopolitas, como o *A. aegypti* e o *A. albopictus* (GAO; SONG; ZHANG, 2019). Tais características, aliadas à aspectos do desenvolvimento socioeconômico dos países tropicais, mudanças climáticas, desmatamento, urbanização sobre áreas silvestres, viagens internacionais, etc., contribuem para o potencial surgimento de novas epidemias de arboviroses e re-emergência de doenças virais infecciosas em todo o planeta (MENDES; KUHN, 2018; RANDOLPH; ROGERS, 2010; RANGEL et al., 2021; SEMENZA; SUK, 2018).

As infecções causadas por alguns destes vírus afetam principalmente as pessoas mais pobres do mundo e são consideradas Doenças Tropicais Negligenciadas (LINDOSO; LINDOSO, 2009). Trata-se de doenças historicamente negligenciadas por financiadores, pesquisadores, políticos e outros tomadores de decisões ao redor do mundo, e predominantemente distribuídas

nos trópicos, onde há elevada concentração de populações pobres, como comunidades rurais remotas e favelas, com baixa assistência básica em saúde, acesso precário à água potável e a saneamento básico (FEASEY et al., 2010; MOLYNEUX, 2013).

2.2 Alphavirus: Estrutura e Características

Os *Alphavirus* são estruturalmente similares, vírion de aproximadamente 70 nm de diâmetro e simetria icosaédrica, e conservam uma configuração genômica próxima entre si (KIM et al., 2021; LEUNG; NG; CHU, 2011). O seu genoma possui 12 kb em tamanho e codifica dois quadros de leitura aberta ou open reading frames (ORF's) cercados por região não traduzidas, nas porções 5' e 3', e separados pelo promotor para o mRNA subgenômico (RYMAN; KLIMSTRA, 2008) (Figura 1). O primeiro quadro de leitura codifica uma poliproteína que é clivada em 4 proteínas não estruturais (nsp1-4) e o segundo codifica uma poliproteína estrutural que é clivada em 5 proteínas estruturais (C, 6K, E1, E2 e E3) (LEUNG; NG; CHU, 2011). As proteínas não-estruturais participam dos processos replicativos da partícula viral, atuando como proteases, no remodelamento de membranas, na fosforilação de moléculas, etc. e as proteínas estruturais participam da montagem da partícula viral e das etapas de atracamento e entrada nas células (RANGEL et al., 2021)(RANGEL et al., 2021; RYMAN; KLIMSTRA, 2008).

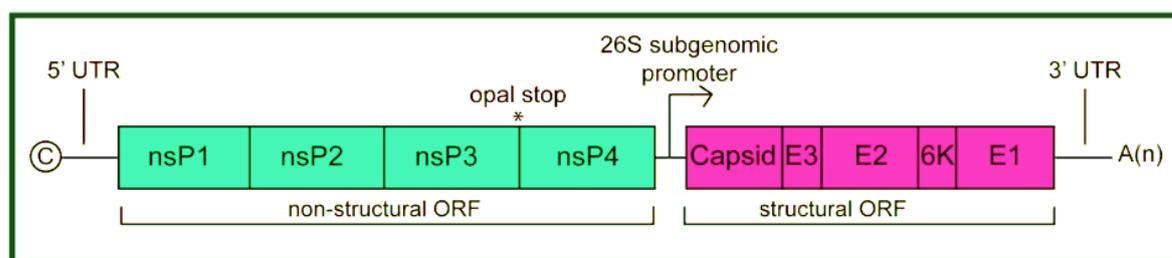


Figura 1: Organização geral do genoma dos Alphavirus (adaptado de RANGEL et al., 2021). Podem ser observados os quadros de leitura aberta ou ORF's que codificam as proteínas não-estruturais (em verde) e estruturais (em rosa), separadas pela porção do promotor subgenômico 26S e margeadas pelas regiões não-traduzidas (UTR's).

Os *Alphavirus* de interesse médico são caracterizados informalmente quanto a localização geográfica de sua primeira identificação em *Alphavirus* do velho mundo, como o *Chikungunya vírus* (CHIKV), *Onyong-nyong virus* (ONNV), *Ross River virus* (RRV), *Semliki Forest virus* (SFV) e *Sindbis virus* (SINV), e *Alphavirus* do novo mundo, que correspondem às espécies *Mayaro virus* (MAYV), *Eastern equine encephalitis virus* (EEEV), *Venezuelan equine*

encephalitis virus (VEEV) e *Western equine encephalitis virus* (WEEV) (LEUNG; NG; CHU, 2011; RYMAN; KLIMSTRA, 2008). Os *Alphavirus* do velho mundo geralmente estão associados à febre, erupções cutâneas, síndromes artríticas graves e outras séries de manifestações reumáticas, e por isso são também conhecidos como *Alphavirus* artritogênicos (KIM et al., 2021; LEUNG; NG; CHU, 2011; NGUYEN et al., 2020). Com exceção do MAYV - filogeneticamente mais relacionado aos *Alphavirus* artritogênicos (Figura 2), os *Alphavirus* do novo mundo são também conhecidos como *Alphavirus* encefalitogênicos por infectarem o sistema nervoso central e causarem doenças que variam de quadros febris leves a encefalites severas, com risco de morte (KIM et al., 2021; STROMBERG et al., 2020).

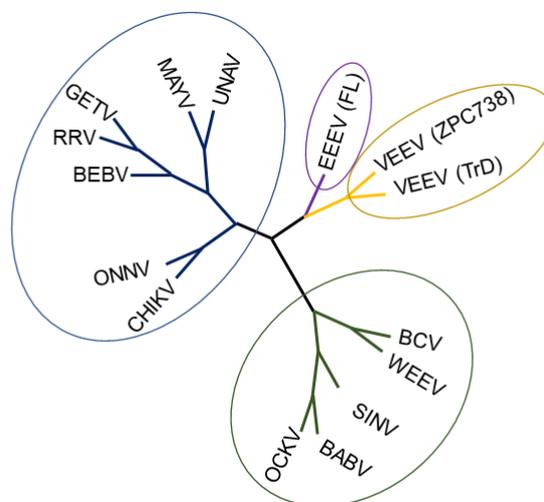


Figura 2: Relações filogenéticas de *Alphavirus* de interesse médico baseadas na sequência de proteínas estruturais E1 e determinada por algoritmos de alinhamento de sequências. Na imagem, os grupos próximos estão representados em grandes complexos: Semliki Forest (círculo azul), WEE (círculo verde), VEE (círculo amarelo) e EEE (círculo roxo) (adaptado de KIM et al., 2021).

2.3 Partícula Viral e Proteínas estruturais

A partícula viral, tipo de nanopartícula encontrada na natureza, diz respeito a estrutura molecular dos vírus e inclui o material genético (DNA ou RNA) e as camadas de proteínas que protege esses genes, como proteínas do capsídeo e em alguns casos, proteínas do envelope (RYU, 2017) (Figura 3).

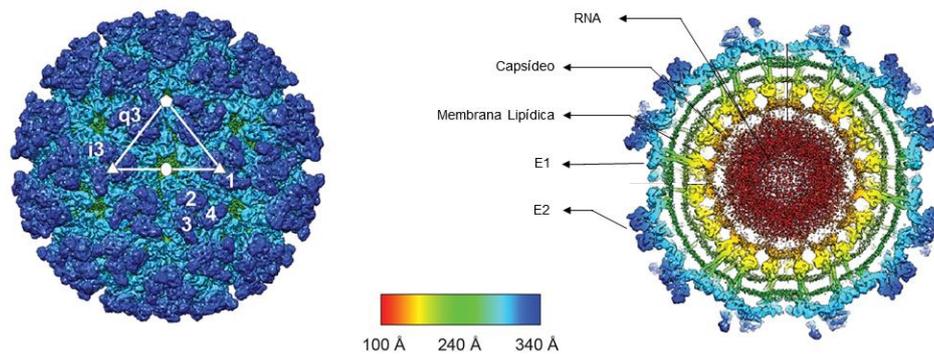


Figura 3: Imagem de cristal de crio-microscopia eletrônica da partícula de um VLP de CHIKV (adaptado de (BASORE et al., 2019)). Estão representados os principais componentes da partícula viral: as proteínas do envelope: E1 (Azul Claro) e E2 (Azul Escuro), membrana lipídica (verde), proteínas do Capsídeo (amarelo) e núcleo de densidade eletrônica e RNA (vermelho).

Nos Alphavirus, as proteínas do capsídeo (ou proteínas C) interagem com o RNA genômico e com as glicoproteínas virais. Desta forma, promovem e/ou auxiliam atividades como encapsulamento do genoma, e assim como a proteína 6K, participa dos processos de montagem e brotamento da partícula viral, além de contribuir para eficiência do estabelecimento da infecção e modulação da resposta inata do sistema imune (RANGEL et al., 2021; RAO; TAYLOR, 2021). As proteínas E1 e E2 de *Alphavirus* são glicoproteínas transmembranares com diversas funções. A proteína E1 é responsável pela ligação com proteínas que medeiam a fusão de membranas durante a infecção e a proteína E2 é responsável, principalmente, pelo atracamento aos receptores celulares e aos fatores de entrada (GAO; SONG; ZHANG, 2019). A proteína E3 possui função de translocação da poliproteína estrutural para o retículo endoplasmático e também na maturação do vírus.

O dímero formado pelas proteínas E1 e E2 é importante para o processamento de proteínas, passagem por membranas e montagem da partícula viral em *Alphavirus*. Estas proteínas possuem ectodomínios importantes para o reconhecimento do sistema imune (HASAN et al., 2018). Os ectodomínios da proteína E2 são divididos em domínios A, B e C (figura 4). Os domínios A e B foram reportados como cruciais para o reconhecimento por anticorpos, para a ligação à célula hospedeira e nas mudanças conformacionais da partícula viral, desencadeadas pelo pH (LI et al., 2010; VOSS et al., 2010). Desta forma, estes são alvos promissores para o estudo e prospecção de alvos antigênicos para intervenções terapêuticas e vacinas.

PROTEÍNAS DO ENVELOPE EM VISTA TRIDIMENSIONAL

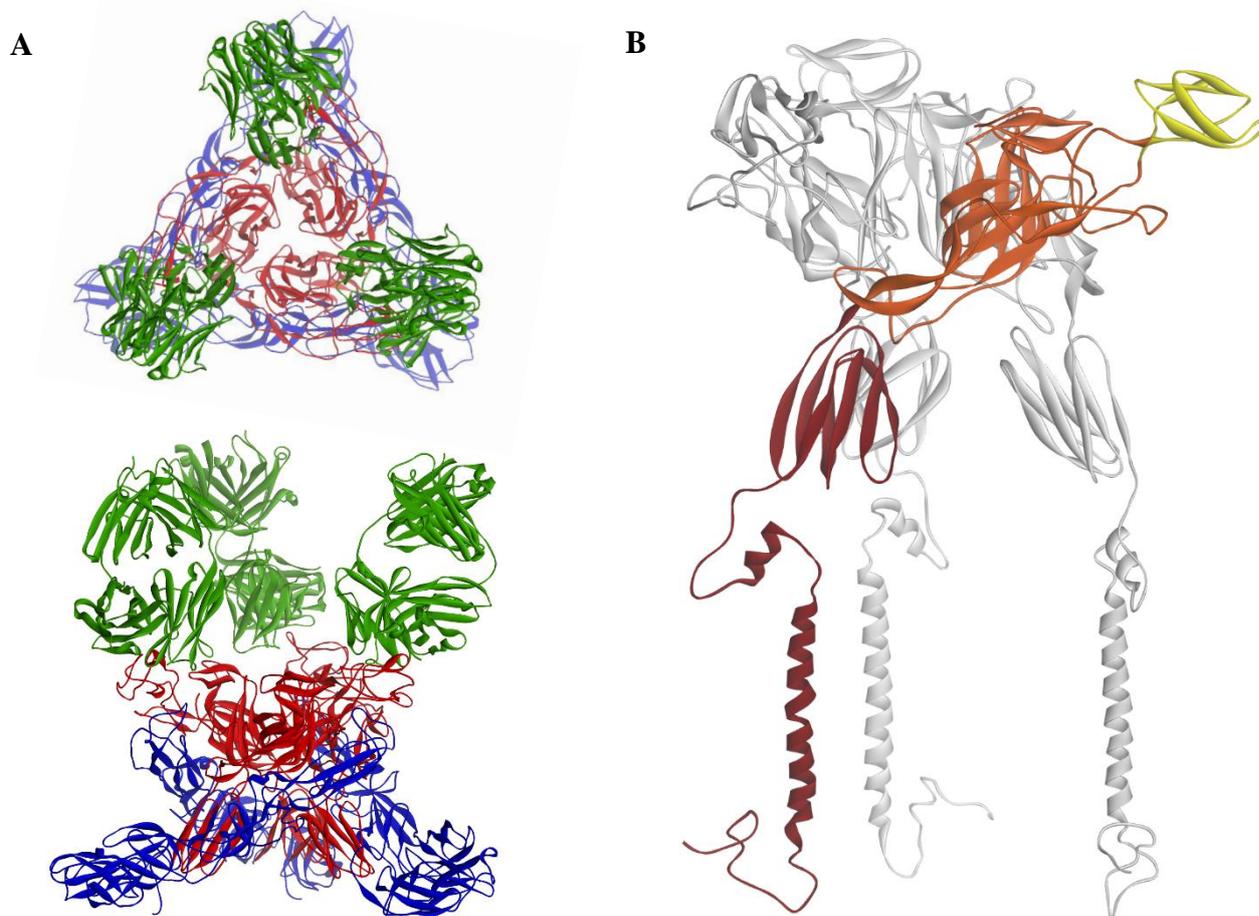


Figura 4: (A) Vistas radial e em rotação 90 ° da estrutura formada pela interação dímero E1-E2 com Fab de anticorpo EEEV-3, onde a proteína E1 está em azul, a proteína E2 em vermelho e as cadeias leve e pesada do anticorpo em verde (Adaptado de HASAN et al. (2018). Acesso no RSCB PDB: 6MW9). (B) Estrutura tridimensional da proteína E2, a partir do cristal de CHIKV (Acesso 6NK7). Estão evidenciados os domínios A (laranja), o domínio B (amarelo) e o Domínio C (vinho). Fonte da imagem: própria

2.4 Peptídeos Imunogênicos

Os epítomos imunogênicos são pequenas porções de um antígeno, como por exemplo cadeias de resíduos de aminoácidos de uma proteína, capazes de se ligarem à receptores antígeno-específicos de linfócitos ou anticorpos e, assim, elicitar uma resposta humoral ou celular (YANG; YU, 2009). Em geral, estes epítomos têm um comprimento mínimo de 5-7 resíduos de aminoácidos e podem ser classificados em epítomos contínuos (ou lineares) e epítomos descontínuos (ou conformacionais) (KANDUC, 2005). Os epítomos contínuos são formados por resíduos de aminoácido distribuídos de forma contígua, ou seja, lado a lado. Os epítomos

descontínuos, por sua vez, são separados na sequência de aminoácidos e trazidos em contiguidade física por dobramento polipeptídico (BUI et al., 2007).

2.5 Nova era das vacinas e vacinologia reversa

Segundo a Organização Mundial de Saúde (OMS), a vacinação é o método mais efetivo no combate a doenças infecciosas (WHO, 2020a). Tradicionalmente, as vacinas são baseadas em etapas de isolamento, inativação e injeção ou inoculação de microrganismos atenuados ou inativados, ou partes destes (L. DUDEK et al., 2010). Trata-se de uma estratégia para preparar as defesas naturais do corpo humano contra infecções específicas à medida que expõe previamente o sistema imune a determinados patógenos, ou partes destes, criando memória e o tornando mais resistente (WHO, 2021).

Nas últimas décadas, os avanços no entendimento da biologia básica de patógenos, nos mecanismos moleculares da resposta imune protetiva humana e das técnicas em biologia molecular, biofísica e informática, entre outras disciplinas, proporcionaram um grande impulso as tecnologias de desenhos de vacinas (MODJARRAD; KOFF, 2016). Atualmente, diferentes estratégias para a construção de vacinas vem sendo desenvolvidas e testadas, como vacina de vírus inativado, vacina de sub-unidade, vacina de vírus atenuado (LAV), vacina de vetor viral recombinante, vacina quimérica, vacina de partícula semelhante a vírus (VLP) e, mais recentemente, vacinas de ácidos nucleicos (GAO; SONG; ZHANG, 2019)), cada uma com seus pontos fortes e fracos.

A partir dos avanços da compreensão dos campos da imunologia humana e da biologia estrutural, novas técnicas, reagentes e ferramentas tem permitido o aprimoramento das estratégias para o desenho de vacinas com base na biologia básica e nos mecanismos moleculares da resposta imune protetiva humana (RAPPUOLI et al., 2016). Neste sentido, as recentes técnicas de sequenciamento de genoma revolucionaram o processo de descobertas de novos antígenos para vacinas através da vacinologia reversa.

A Vacinologia reversa é o processo de descoberta de antígenos com maior probabilidade de serem candidatos a vacinas a partir das informações genômicas dos patógenos (RAPPUOLI, 2000). Esta abordagem vem sendo amplamente utilizada nas últimas décadas e configura um

método bem-sucedido de descoberta de vacinas, sendo passível de ser explorada contra muitos tipos de patógenos, sem necessariamente ter que produzi-los em laboratório.

Utilizando as ferramentas computacionais e grandes conjuntos de dados biológicos, como dados ômicos e bancos de dados de imunologia, reduziu-se os custos e o tempo necessário para construção de vacinas à medida que permitiu a seleção de uma lista reduzida e otimizada de antígenos candidatos aptos a serem testados em modelos animais (DEL TORDELLO; RAPPUOLI; DELANY, 2017).

Atualmente, as abordagens de vacinologia reversa incluem análises comparativas *in silico* entre múltiplas sequências genômicas, permitindo a identificação de antígenos conservados dentro de uma população heterogênea de patógenos, ou ainda, a identificação de antígenos únicos, ou seja, que são exclusivos de isolados patogênicos (DEL TORDELLO; RAPPUOLI; DELANY, 2017).

2.6 Abordagens *in silico* e Imunoinformática

As abordagens *in silico* compreendem um conjunto de técnicas que utilizam modelos computacionais para simular/investigar sistemas biológicos. Com o advento da bioinformática, estas técnicas tem sido empregadas para auxiliar atividades laboratoriais a atingir resultados assertivos, com maiores índices de confiabilidade, economizando recursos financeiros e tempo (YANG; BOGDAN; NAZARIAN, 2021).

Nas últimas décadas, a aplicação dos avanços nas ferramentas computacionais sobre grandes conjuntos de dados experimentais de imunologia possibilitou a criação de um novo campo de estudo, a imunoinformática (FLOREA et al., 2003; YANG; YU, 2009). Esta área do conhecimento auxilia os cientistas a compreender cada vez mais o sistema imune, permitindo a construção de modelos e análises mais assertivos e robustos sobre os dados do sistema imune humano e patógenos infecciosos.

Como forma de otimizar os trabalhos de construção de candidatos vacinais, as técnicas de imunoinformática tem sido utilizadas na busca de alvos para a construção de testes diagnósticos e vacinas, através da predição e seleção de epítomos antigênicos (YANG; YU, 2009). As

técnicas empregadas nestas análises consistem na aplicação de algoritmos validados e de amplo conhecimento na determinação de alvos da resposta imunológica, sua estrutura e biologia (BUI et al., 2007; GUPTA R.; JUNG E.; BRUNAK S., 2004; JESPERSEN et al., 2017; KOLASKAR; TONGAONKAR, 1990; KROGH et al., 2001; STEENTOFT et al., 2013). Os algoritmos de predição de epítomos reduzem tempo e custos associados à prospecção de alvos antigênicos, uma vez que reduzem a quantidade de testes em bancada, e auxiliam a construção de vacinas e testes diagnósticos seguros, eficazes e acessíveis.

2.7 Predições de epítomos

A identificação de epítomos em antígenos virais e não-virais é importante para compreensão da etiologia da doença, monitoramento imunológico, desenvolvimento de ensaios de diagnóstico e desenvolvimento de vacinas baseadas em epítomos (SANCHEZ-TRINCADO; GOMEZ-PEROSANZ; RECHE, 2017). No entanto, a identificação tradicional destes epítomos frequentemente exige trabalhos dispendiosos, como a resolução da estrutura tridimensional de complexos antígeno-anticorpo e triagem de bibliotecas de peptídeos de ligação de anticorpos (epítomos de células B), e experimentos em multímeros de MHC e ensaios de linfoproliferação ou ELISPOT (epítomos de células T) (SANCHEZ-TRINCADO; GOMEZ-PEROSANZ; RECHE, 2017).

Neste sentido, um grande avanço para este campo foi desenvolvimento de algoritmos de previsão *in silico* de epítomos, capazes de reduzir a carga associada ao mapeamento destes epítomos, diminuindo a lista de potenciais candidatos a serem testados experimentalmente. Esta expansão possibilitou, por exemplo, a identificação de alvos antigênicos de forma rápida e menos laboriosa e tem permitido a simplificação de problemas reais de imunologia e a construção de modelos experimentais mais eficazes e competentes (LARSEN et al., 2007; YANG; YU, 2009). Atualmente, as abordagens *in silico* para predição de epítomos estão distribuídas em quatro grandes métodos: métodos baseados em sequência, métodos baseados em estrutura, métodos híbridos e métodos baseados em consenso. Segundo YANG & YU (2009), os métodos baseados em consenso são considerados os melhores de predição, seguidos dos métodos híbridos, que por sua vez, tem acurácia maior que as abordagens de métodos baseados em sequência e estrutura.

3. OBJETIVOS

3.1 Objetivo geral:

- Predizer através de abordagens *in silico* peptídeos imunogênicos conservados entre CHIKV e EEEV que possibilitem a construção de um protótipo vacinal para estes dois *Alphavirus*.

3.2 Objetivos específicos:

- Predizer peptídeos imunogênicos conservados dentre as proteínas estruturais E2 dos vírus do gênero *Alphavirus*: CHIKV e EEEV;
- Discriminar epítomos de proteínas estruturais dos *Alphavirus* quanto às características físico-químicas, posição, acessibilidade ao solvente, antigenicidade, alergenicidade e toxicidade;
- Mapear a região dos epítomos nas proteínas virais alvo;
- Apontar bons alvos putativos para construção de modelos vacinais para CHIKV e EEEV.

4. MATERIAIS E MÉTODOS

4.1 Aquisição de Sequências

Para a predição de epítomos foram utilizadas as sequências das proteínas E2 obtidas a partir das sequências de referência das poliproteínas estruturais dos Alphavirus de interesse médico: o vírus artrítogênico CHIKV e o vírus encefalitogênico EEEV. Estas sequências estão disponíveis no GenBank® (<https://www.ncbi.nlm.nih.gov/genbank/>) sob os identificadores NP_690589.2 e NP_632022.1, respectivamente. As poliproteínas estruturais, também utilizadas na análise de conservação, foram derivadas dos genomas de referência destes vírus (KHAN et al., 2002; VOLCHKOV; VOLCHKOVA; NETESOV, 1991).

A glicoproteína E2 do CHIKV apresenta tamanho de 423 resíduos de aminoácidos, localizados na posição 326-748 da poliproteína estrutural. A glicoproteína E2 do EEEV apresenta tamanho de 420 resíduos de aminoácidos e está localizada na posição 325-744 da poliproteína estrutural. Ambas apresentando grande porção extracelular como domínio na topologia de membrana.

As sequências proteicas (Anexo 8.1) foram copiadas e submetidas à ferramenta Sequence Manipulation Suite: Filter Protein (STOTHARD, 2000) para validação da integridade e avaliação da presença de resíduos não resolvidos ('x'). As sequências completas das poliproteínas estruturais completas e das proteínas E2 foram armazenadas em um banco de dados próprio, utilizando-se arquivos fasta.

4.2 Predição de epítomos

As predições de epítomos de células B, células T-citotóxicas e células T-auxiliares foram realizadas em ferramentas web de predição de epítomos amplamente utilizadas na literatura (PETERS; NIELSEN; SETTE, 2020; SORIA-GUERRA et al., 2015). Cada uma destas ferramentas utilizam técnicas computacionais distintas para obter uma boa performance na identificação da sequência de antígenos (SANCHEZ-TRINCADO; GOMEZ-PEROSANZ; RECHE, 2017).

Para a predição de epítomos de células B, as sequências de referências das proteínas E2 foram submetidas às ferramentas web de predição de epítomos lineares de células B: Bepipred 2.0 (<https://services.healthtech.dtu.dk/service.php?BepiPred-2.0>) (JESPERSEN et al., 2017) e Predicted Antigenic Peptides/IMED (PAP/IMED) (<http://imed.med.ucm.es/Tools/antigenic.pl>) (KOLASKAR; TONGAONKAR, 1990).

Para a predição no Bepipred 2.0 foram utilizados o valor padrão (default) para o parâmetro de seleção de epítomos (“Epitope Threshold” = 0,5). Os resultados das predições foram baixados em um único arquivo JSON, disponibilizado pelo método, contendo as predições para todas as proteínas. A ferramenta PAP/IMED, por sua vez, não oferece parâmetros a serem ajustados e a predição ocorreu via submissão de uma única sequência proteica por vez. Os resultados, uma tabela em HTML para cada uma destas predições, foram copiados e salvos em um único arquivo, formato TXT (Anexo 8.2).

Para a predição de epítomos de células T-citotóxicas, as sequências de referências das proteínas E2 foram submetidas à ferramenta web de predição NetCTL 1.2 (<https://services.healthtech.dtu.dk/service.php?NetCTL-1.2>) (LARSEN et al., 2007). Os valores de parâmetros utilizados seguem o padrão indicado pelo método: O supertipo de HLA selecionado foi o A1, para o peso relativo da clivagem proteassômica na porção C terminal foi utilizado o valor 0,15, para o peso relativo da eficiência de transporte da TAP foi utilizado o valor de 0,05 e como limiar de identificação dos epítomos adotou-se o valor 0,75. Não foi selecionada a opção de organização dos epítomos na tabela de resultados em ordem decrescente dos valores de pontuação atribuídos pelo método. As tabelas com os respectivos epítomos e as pontuações atribuídas pelo método foram baixadas e salvas em um arquivo de formato HTML.

Para a predição de epítomos de células T-auxiliares, as sequências de referências das proteínas E2 foram submetidas à ferramenta web de predição [MHC-II Binding Predictions](http://tools.iedb.org/mhcii/) (<http://tools.iedb.org/mhcii/>). O método de predição selecionado foi o “IEDB recommended 2.22”. Foram selecionados, também: o locus para HLA DR humano utilizando-se o conjunto de alelos referência para HLA, disponibilizados pela própria ferramenta; O tamanho padrão de 15 aminoácidos para as sequências dos peptídeos; A organização do resultado final segundo o ranqueamento ajustado oferecido pelo método, baseado nas frequências dos comprimentos dos peptídeos, e a apresentação deste resultado como arquivo de texto (“text file”). Os resultados

das predições de células T-auxiliares foram baixados como arquivos HTML, um para cada uma das proteínas E2, uma vez que a predição recebe uma única proteína por vez. Ambos os arquivos foram salvos e armazenados em um mesmo diretório.

4.3 Triagem dos resultados das predições de epítomos

Os resultados das predições de epítomos foram selecionados computacionalmente através da ferramenta POA1 (https://github.com/UbiratanBatista/POA_Project), que corresponde a primeira metade da triagem otimizada de epítomos do pipeline de otimização de antígenos (POA). O POA refere-se a um conjunto de passos dentro de uma estratégia de construção in silico de modelos vacinais. Para execução deste pipeline são utilizadas múltiplas ferramentas computacionais, como aquelas de triagem otimizada utilizadas neste trabalho (primeira fase do pipeline).

As ferramentas utilizadas neste trabalho foram desenvolvidas em Python 3.8 com o intuito de facilitar a seleção de sequência e informações associadas a epítomos dentre o volume de dados que é retornado das ferramentas web de predição de epítomos. Estas ferramentas foram implementadas via terminal em sistema Linux (Ubuntu 20.04 LTS).

Os arquivos resultantes da predição de epítomos: JSON e TXT (predições de células B), HTML (predição de células T - citotóxicas) e o diretório com os arquivos HTML (predições de células T – auxiliares), foram utilizados como dados de entradas para a ferramenta POA1 (parâmetros -b, -p, -n e -m, respectivamente) (Figura 5). A ferramenta recebeu também o arquivo fasta contendo as sequências das proteínas usadas nas predições (parâmetro -f) e o caminho (ou path) para o diretório onde os resultados foram salvos (parâmetro -d).

No POA1, para a triagem dos epítomos de linfócitos T-auxiliares, foi ajustado o parâmetro do valor do “threshold” de alta afinidade de ligação do epítomo com o MHCII ($IC_{50} < 50$ nM) (parâmetro -mic). Para a seleção dos epítomos de linfócitos B, foram ajustados os parâmetros para o tamanho mínimo de sequência dos epítomos, sendo aceito o valor de 5 (parâmetros -bmin e -pmin). Para dados de saída, o parâmetro E foi modificado (parâmetro -e = y) para que os resultados fossem também organizados em planilha no formato .xlsx.

Com os parâmetros ajustados, a ferramenta POA1 selecionou as sequências dos peptídeos imunogênicos e organizou estes resultados em arquivos fastas, um para cada espécie dentre os vírus analisados, e em uma planilha excel (arquivo .xlsx) (Figura 5). Cada arquivo do tipo fasta recebeu as sequências dos peptídeos de um mesmo vírus e os cabeçalhos de identificação contendo: nome da proteína, método de predição do peptídeo, espécie do organismo, posição inicial e posição final da sequência do peptídeo na proteína (Anexo 8.3).

4.4 Análise de Conservação

Os arquivos fastas contendo os epítomos selecionados pelo POA1, separados pelo organismo de origem, foram submetidos à análise de conservação na ferramenta web do IEDB, [Epitope Conservancy Analysis](http://tools.iedb.org/conservancy/) (<http://tools.iedb.org/conservancy/>) (BUI et al., 2007), junto aos arquivos das poliproteínas estruturais de referência dos vírus. As análises de conservação das sequências de epítomos lineares ocorreram sob um threshold de identidade de sequência $\geq 60\%$ e os resultados foram baixados e salvos como planilhas do tipo CSV. As análises de conservação objetivaram determinar o grau de conservação do epítomo em uma dada sequência proteica definida em um determinado nível de identidade. Desta forma, a análise buscou através de alinhamento o “match” entre as sequências dos peptídeos das proteínas E2 dos Alphavirus e a sequência completa da poliproteína estrutural de referência da outra espécie presente na análise. A triagem não ocorreu entre os peptídeos e a poliproteína estrutural de uma mesma espécie como forma de evitar resultados redundantes.

4.5 Triagem dos resultados da análise de conservação e análise de propensão a regiões trans-membranares

Os arquivos CSV resultantes da análise de conservação foram armazenados em um único diretório destinado aos resultados da ferramenta web Epitope Conservancy Analysis. O diretório contendo estes arquivos foi utilizado como parâmetro de entrada (parâmetro -d) para a ferramenta POA2, que corresponde a segunda metade da triagem otimizada de epítomos do POA (Figura 5).

A ferramenta POA2 é utilizada para a análise dos resultados das predições de conservação e de topologia de membrana. Assim como os arquivos CSV, a POA2 recebeu como argumentos

obrigatórios a natureza da análise (parâmetros `-g = True`) – para encontrar epítomos conservados (\geq), o “threshold” de identidade de sequência avaliado (parâmetro `-t`) igual a 60, o arquivo fasta contendo as poliproteínas estruturais de referência dos vírus presentes na análise (parâmetro `-f`) e o caminho do diretório em que os resultados da análise foram salvos (parâmetro `-r`). Os demais parâmetros seguiram os valores padrões para a ferramenta: identidade mínima de 60%, identidade máxima de 100% e porcentagem de 60 % da sequência da proteína que corresponde a um nível de identidade especificado.

A ferramenta POA2 selecionou e recuperou os epítomos que tiveram mais de 60% de identidade entre os vírus de interesse com base nos scores de identidades definidos pela ferramenta do IEDB. Além disso, utilizando o módulo pyTMHMM, baseado em modelos de Hidden Markov do algoritmo TMHMM (<https://github.com/bosborne/pyTMHMM>) (SONNHAMMER; VON HEIJNE; KROGH, 1998), realizou a predição de hélices transmembranares e classificou os epítomos quanto as topologias de membrana.

Por fim, a ferramenta POA2 retornou uma planilha Excel (.xlsx), contendo os epítomos conservados e a caracterização da topologia de membrana de cada um deles, e um arquivo fasta (parâmetro `-rf = 1`) contendo apenas os epítomos localizados em porções externas da topologia de membrana. Este resultado corrobora um conjunto de sequências preditamente antigênicas, conservadas dentre o grupo de vírus e de topologia conhecida nas proteínas virais.

4.6 Análise de Toxicidade e Alergenicidade

Os epítomos selecionados pelas ferramentas da triagem otimizada de antígenos (POA) foram submetidos a análise de toxicidade e alergenidade através das ferramentas web de predição Toxinpred (<https://webs.iiitd.edu.in/raghava/toxinpred/index.html>) e Allercatpro (<https://allercatpro.bii.a-star.edu.sg/>) (GUPTA et al., 2013; MAURER-STROH et al., 2019), respectivamente. Os resultados de ambos os métodos foram salvos em arquivos de planilha Excel (.xlsx) e nenhum dos peptídeos foi caracterizado como potencialmente tóxico ou alergênico.

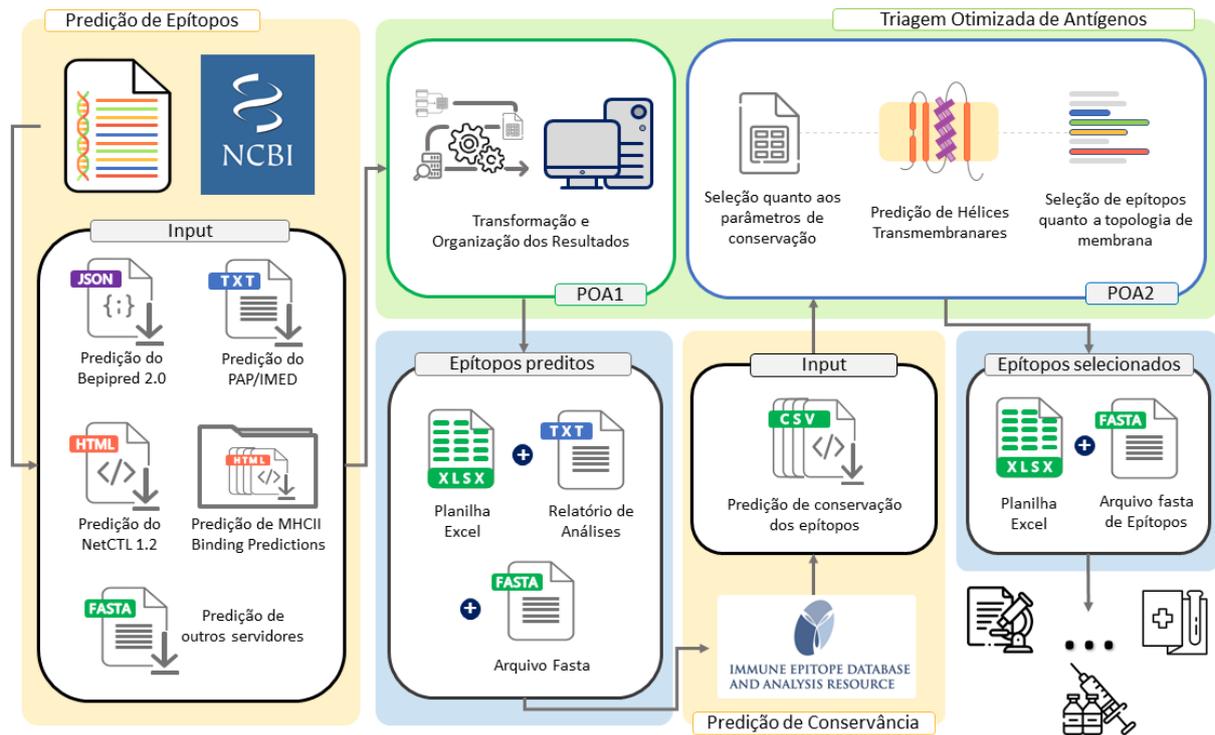


Figura 5: Fluxograma das etapas da Triagem Otimizada de Antígenos. Nos quadros amarelos ocorrem etapas manuais, como submissão de sequências à servidores web e download dos resultados, em verde são as ferramentas computacionais para triagem otimizada de antígenos, POA1 e POA2, e em azul, os resultados de saída das ferramentas de triagem.

4.7 Caracterização Tridimensional dos Alvos

Para obtenção de dados de acessibilidade ao solvente, bem como informações complementares de características físico-químicas, os epítomos conservados triados pelas ferramentas de triagem otimizada de antígenos (POA) foram confrontados em estruturas cristalinas das proteínas E2 dos Alphavirus CHIKV e EEEV. A visualização tridimensional da estrutura destas proteínas foram baixadas a partir de arquivos no formato PDB do Banco de Dados RSCB Protein Data Bank (<https://www.rcsb.org/>), acesso 6NK7 (BASORE et al., 2019) (CHIKV) e 6MX4 (HASAN et al., 2018) (EEEV).

O modelo estrutural da partícula viral de CHIKV (6NK7) possui resolução de 4,99 Å, tendo sido expresso em células Vero de primatas *Chlorocebus aethiops*. O cristal da partícula viral de EEEV (6MX4) possui resolução de 4,40 Å e foi expresso em células BHK-15 de *Mesocricetus auratus*. Ambas os cristais foram avaliados e validados pelas ferramentas MolProbity (<http://molprobity.biochem.duke.edu/index.php?MolProbSID=j6nqm0nrmgjrmpik3j6ip4hq0&eventID=2>) (WILLIAMS et al., 2018) através dos valores do Ramachandran plot, Clashscore,

MolProbity score, e ProSA-web (<https://prosa.services.came.sbg.ac.at/prosa.php>) (WIEDERSTEIN; SIPPL, 2007), através do valor de Z-score.

O Ramachandran plot foi utilizado a fim de determinar as torções permitidas aos ângulos phi (ϕ) e psi (ψ) na estrutura proteica, de acordo as ligações estabelecidas entre os resíduos de aminoácidos. Os valores de Clashscore e MolProbity score, por sua vez, são derivados de uma série de análises para confirmação e validação das estruturas 3D de proteínas, como a análise de sobreposições estéricas entre ligações átomo/átomo (ligações de hidrogênio e do tipo van der Waals), avaliação das geometrias de ligação, etc. Desta forma, o MolProbity combina o clashscore, rotameros e avaliações do Ramachandran em um único score, normalizado para a escala de resolução de raio-X.

O Z-score utilizado pelo ProSA-web é uma pontuação que indica a qualidade geral do modelo tridimensional. Este parâmetro é calculado para cada estrutura, mas é avaliado em um gráfico que contém os z-scores de todas as cadeias proteicas experimentalmente determinadas até o momento no banco de dados de PDB. Desta forma, este parâmetro é utilizado para verificar se o z-score da estrutura analisada está dentro das faixas de pontuação de z-scores tipicamente encontrados para proteínas nativas de tamanho semelhante.

4.8 Cálculo de acessibilidade ao Solvente

Os epítomos selecionados foram identificados nos modelos validados das estruturas tridimensionais das proteínas E2 através do software BIOVIA Discovery Studio 2019 ® (<https://www.3ds.com/products-services/biovia/products/molecular-modeling-simulation/biovia-discovery-studio/>). Através dos modelos da proteína E2, e selecionando apenas uma cadeia do trímero da estrutura da proteína, foram calculados os valores de acessibilidade ao solvente dos peptídeos triados através do software Pymol (<https://pymol.org/2/>). Para o cálculo foi utilizado a fórmula (adaptada de Pymol tools):

$$A_{ac} = \frac{(A_{ep} + A_{pt}) - A_{total}}{2} - A_{ep}$$

Onde Aac é o valor (em Angstrom²) da área exposta ou acessível da superfície do epítopo, Aep é a área da superfície do epítopo separado da proteína, Apt é a área da superfície da proteína tendo sido retirado o epítopo e $Atotal$ refere-se a área de superfície da proteína nativa, sem que tenha sido feitas alterações em sua estrutura.

5. RESULTADOS E DISCUSSÕES

A seleção de alvos antigênicos por meio da triagem otimizada de epítomos do Pipeline de Otimização de Antígenos (POA) promoveu a substituição de etapas anteriormente realizadas manualmente pela seleção e organização de algoritmos computacionais. Esta mudança reflete uma otimização significativa para etapa de triagem e recuperação de informações da predição de epítomos, da predição de conservação e da predição de hélices transmembranares e topologia de membrana, devido a diminuição do uso da inspeção visual e da intervenção humana na captura dos resultados (Figura 6).

Etapas manuais, de uma forma geral, são consideradas abaixo de 100% de eficácia, aumenta à propensão de indução ao erro, dificulta a reprodutibilidade dos ensaios e diminui a produtividade e confiabilidade do resultado final (FISCH, 1998; MEIJERING et al., 2009; PINHEIRO, 2007). Desta forma, os algoritmos do POA utilizados buscaram selecionar os melhores peptídeos pelos scores oferecidos por cada método em cada predição e então, diminuir o tempo e os esforços gastos pelo operador.

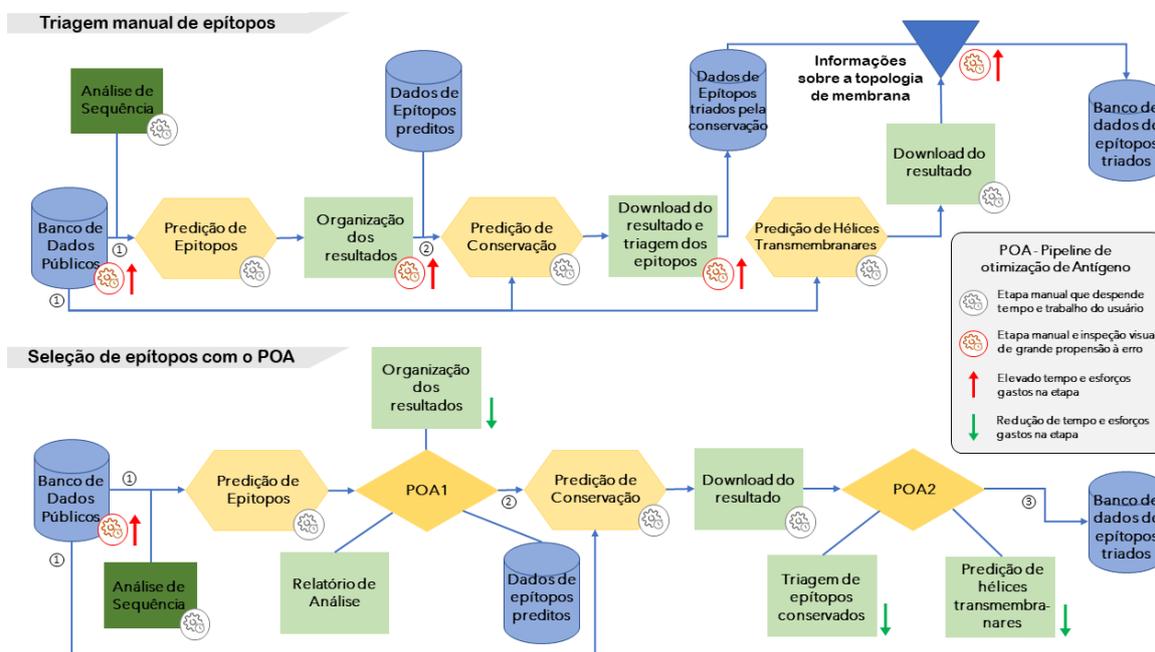


Figura 6: Comparação entre as etapas da seleção de epítomos realizadas manualmente e com o uso da triagem otimizada do POA. Etapas que envolvam coleta/seleção de dados (sequências, ID de proteínas, etc.) e edição de arquivos necessitam de maior cuidado e atenção, uma vez que erros nestas etapas podem prejudicar o resultado final da análise. Neste sentido, as ferramentas de triagem otimizada do POA reduzem a necessidade da intervenção manual nestas etapas (indicado pelas setas verde) e, conseqüentemente, a ocorrência de etapas muito propensas à erros (indicado pela engrenagem em vermelho).

A análise do POA1 para os resultados de predição de epítomos das proteínas E2 de CHIKV e EEEV identificou um total de 294 peptídeos preditivamente antigênicos. O maior percentual dos epítomos preditos são para células T-auxiliares (71,20%), seguido de epítomos para células B (19,92%) e para células T-citotóxicas (8,87%) (Tabela 1). Por sua vez, a análise dos resultados da predição de conservação e da predição de hélices transmembranares realizada pelo POA2 selecionou 63 epítomos destes peptídeos (21.50%) como sendo conservados dentre as proteínas E2 dos Alphavirus analisados, com 60% de identidade mínima entre as sequências.

As diferenças e similaridades entre as proteínas E2 destes vírus foram verificadas através de alinhamento global para-a-par apontou identidade de 41,3% e similaridade de 56,2% entre as sequências (Anexo 8.4). A representação da conservação entre os epítomos preditos e as sequências das proteínas E2 foi também verificada em um alinhamento múltiplo de sequências (Anexo 8.5). Este alinhamento foi realizado utilizando o algoritmo ClustalW (THOMPSON; HIGGINS; GIBSON, 1994), implementado no software MEGA-X (KUMAR et al., 2018).

Tabela 1: Quantidades de epítomos da proteína E2 dos Alphavirus CHIKV e EEEV preditos pelos métodos de predição de antígenos: Bepipred (Linfócitos B), PAP/IMED (Linfócitos B), NetCTL (Linfócitos T – Citotóxicos) e MHCII Binding Predictions (Linfócitos T – auxiliares) e selecionados pela análise da ferramenta POA1.

Espécie	Bepipred	IMED	NetCTL	MHCIIBD	Total
<i>Chikungunya virus</i>	11	19	11	88	129
<i>Eastern equine encephalitis virus</i>	13	13	16	123	165
Total					294

O alinhamento múltiplo de sequencias apontou três motivos conservados promissores para candidatos vacinais, uma vez que ambos os vírus tiveram epítomos preditos para a região. Alguns trabalhos (KROON CAMPOS et al., 2020; LI et al., 2011; NGUYEN et al., 2020) demonstraram que a conservação de motivos antigênicos é interessante do ponto de vista de abordagem vacinais pois podem, até certo ponto, oferecer imunidade inespecífica para mais de um espécie viral, tanto para Alphavirus quanto para arbovírus filogeneticamente correlatados.

Curiosamente, apenas um dos motivos conservados com epítomos preditos para os dois vírus está localizado no domínio A da proteína E2, os outros dois estão localizados no domínio C. Os domínios A e B da proteína E2 são conhecidos por interagirem com regiões Fabs de anticorpos (HASAN et al., 2018). No entanto, os epítomos preditos para E2 de CHIKV e EEEV triados na análise e pertencentes à esta porção da proteína são epítomos para células T-auxiliares.

Alguns dos epítomos preditos para células T- auxiliares foram redundantes, pois foram preditos para mais de um tipo de alelo de HLA. Desconsiderando os epítomos redundantes (Anexo 8.6), 40 (63,5%) dos 63 epítomos anteriormente selecionados foram contabilizados como sendo conservados dentre os Alphavirus CHIKV e EEEV (Tabela 3). Através do módulo pyTMHMM, o POA2 classificou estes epítomos quanto à sua posição na membrana e concluiu que 30 (75%) destes epítomos estão em porções expostas, 10 (25%) em regiões internas, e nenhum (0%) em regiões transmembranares da proteína E2.

A presença de sítios conservados nas regiões acessíveis da proteína E2, responsáveis pela interface de ligação com receptores celulares podem estar relacionados à manutenção de vias de reconhecimento e entrada celular dentre estes vírus (BYRD; KIELIAN, 2017). Estas porções podem ser exploradas para a construção de plataformas vacinais *Pan-Alphavirus*.

Tabela 2: Quantidade de epítomos conservados entre as proteínas E2 dos Alphavirus CHIKV e EEEV separados pelos métodos de predição de epitopos avaliados. Estes epítomos são resultados da análise da ferramenta POA2.

Espécie	Bepipred	IMED	NetCTL	MHCIIBD	Total
<i>Chikungunya virus</i>	1	6	0	14	21
<i>Eastern equine encephalitis virus</i>	1	1	2	15	19
Total					40

As proteínas E2 são os principais alvos da resposta de anticorpos neutralizantes, o que pode ser associado a exposição dos seus grandes ectodomínios A e B, devido sua ligação da partícula viral à glicosaminoglicanos e receptores da superfície celular, assim como na entrada da partícula viral (VANEY; DUQUERROY; REY, 2013).

Tabela 3: Epitopos conservados das proteínas E2 de CHIKV e EEEV, triados com base nas ferramentas de predição de epítipo, de análise de conservação e de predição de porções transmembranares e selecionados pelas ferramentas POA 1 e POA2.

Espécie	Método de Predição	Posição Inicial	Posição Final	Sequência	Comprimento da Sequência	Identidade entre Sequências (%)	Topologia Externa	Topologia Interna
CHIKV	Bepipred	207	212	NEGLIT	6	66,67	x	
CHIKV	PAP/IMED	80	92	RAGLFVRTSAPCT	13	61,54	x	
CHIKV	PAP/IMED	98	105	GHFILARC	8	62,5	x	
CHIKV	PAP/IMED	109	115	ETLTVGF	7	71,43	x	
CHIKV	PAP/IMED	237	244	YNSPLVPR	8	62,5	x	
CHIKV	PAP/IMED	282	297	QVIMLLYPDHPHTLLSY	16	62,5	x	
CHIKV	PAP/IMED	404	419	PGATVPFLLSLICCIR	16	62,5		x
CHIKV	MHCII-Binding	6	20	FNVYKATRPYLAHCP	15	60	x	
CHIKV	MHCII-Binding	8	22	VYKATRPYLAHCPDC	15	60	x	
CHIKV	MHCII-Binding	98	112	GHFILARCPKGETLT	15	60	x	
CHIKV	MHCII-Binding	99	113	HFILARCPKGETLTV	15	60	x	
CHIKV	MHCII-Binding	280	294	KNQVIMLLYPDHPHTL	15	60	x	
CHIKV	MHCII-Binding	281	295	NQVIMLLYPDHPHTLL	15	66,67	x	
CHIKV	MHCII-Binding	282	296	QVIMLLYPDHPHTLLS	15	66,67	x	
CHIKV	MHCII-Binding	283	297	VIMLLYPDHPHTLLSY	15	66,67	x	
CHIKV	MHCII-Binding	289	303	PDHPHTLLSYRSMGEE	15	66,67	x	
CHIKV	MHCII-Binding	290	304	DHPHTLLSYRSMGEEP	15	60	x	
CHIKV	MHCII-Binding	291	305	HHPHTLLSYRSMGEEP	15	60	x	
CHIKV	MHCII-Binding	397	411	ITPYELTPGATVPFL	15	66,67		x
CHIKV	MHCII-Binding	398	412	TPYELTPGATVPFLL	15	66,67		x
CHIKV	MHCII-Binding	399	413	PYELTPGATVPFLLS	15	60		x

Tabela 3: Epitopos conservados das proteínas E2 de CHIKV e EEEV, triados com base nas ferramentas de predição de epítipo, de análise de conservação e de predição de porções transmembranares e selecionados pelas ferramentas POA 1 e POA2 (Continuação).

Espécie	Método de Predição	Posição Inicial	Posição Final	Sequência	Comprimento da Sequência	Identidade entre Sequências (%)	Topologia Externa	Topologia Interna
EEEV	Bepipred	323	332	LEYTWGNHPP	10	70	x	
EEEV	PAP/IMED	142	152	PEHGVLPENR	11	63,64	x	
EEEV	NetCTL	145	153	GVELPCNRY	9	66,67	x	
EEEV	NetCTL	317	325	TVTGELEY	9	66,67	x	
EEEV	MHCII-Binding	6	20	FTQYKLARPYIADCP	15	60	x	
EEEV	MHCII-Binding	8	22	QYKLARPYIADCPNC	15	60	x	
EEEV	MHCII-Binding	9	23	YKLARPYIADCPNCG	15	66,67	x	
EEEV	MHCII-Binding	174	188	DHLLSIHSAKVKIT	15	60	x	
EEEV	MHCII-Binding	175	189	HLLSIHSAKVKITV	15	60	x	
EEEV	MHCII-Binding	176	190	LLSIHSAKVKITVP	15	60	x	
EEEV	MHCII-Binding	177	191	LLSIHSAKVKITVPS	15	66,67	x	
EEEV	MHCII-Binding	277	291	HRTLILHLYPDHPTL	15	60	x	
EEEV	MHCII-Binding	278	292	RTLILHLYPDHPTLL	15	66,67	x	
EEEV	MHCII-Binding	392	406	LCITPYKLAPNAQVP	15	66,67		x
EEEV	MHCII-Binding	393	407	CITPYKLAPNAQVPI	15	66,67		x
EEEV	MHCII-Binding	394	408	ITPYKLAPNAQVPIL	15	66,67		x
EEEV	MHCII-Binding	395	409	TPYKLAPNAQVPILL	15	66,67		x
EEEV	MHCII-Binding	396	410	PYKLAPNAQVPILLA	15	60		x
EEEV	MHCII-Binding	397	411	YKLAPNAQVPILLAL	15	60		x

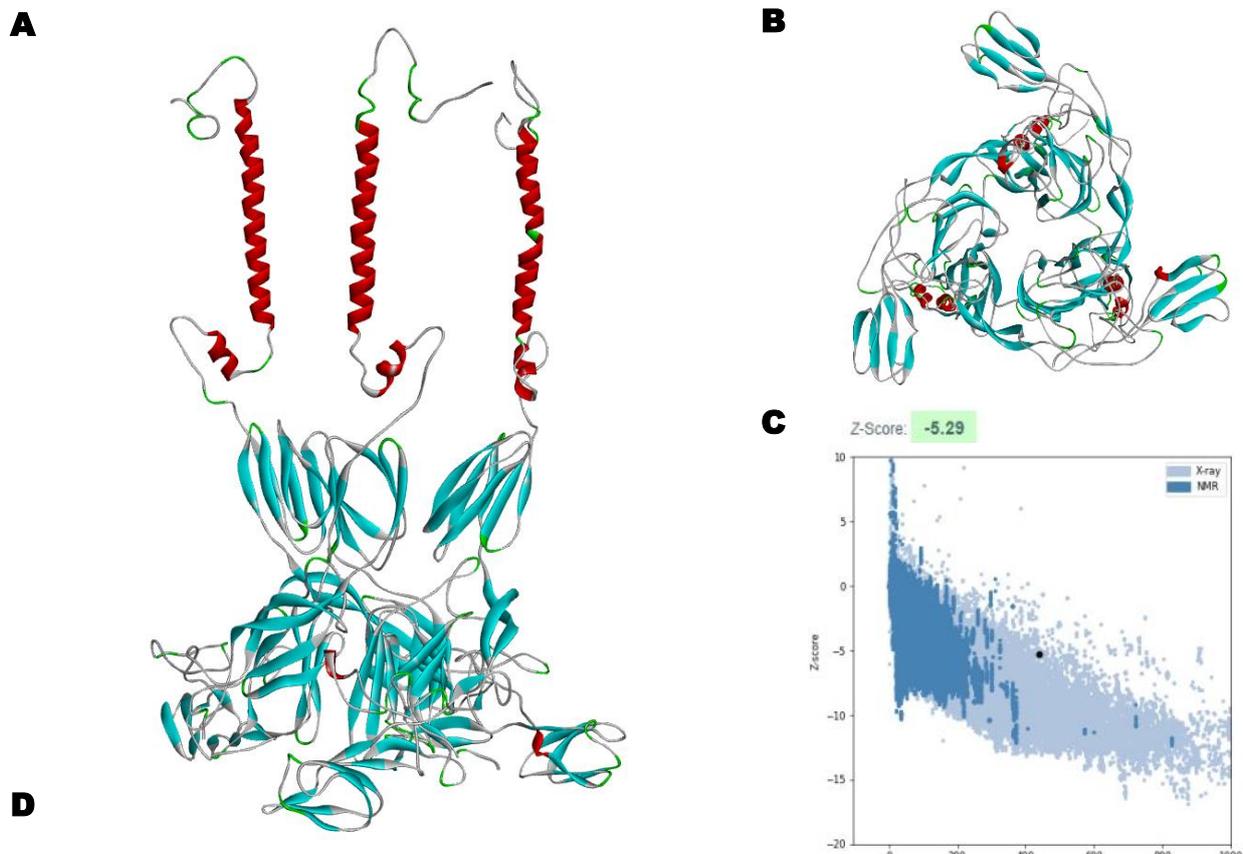
Os epítomos selecionados pelo POA2 foram confrontados com as estruturas dos cristais de criomicroscopia eletrônica das proteínas E2 de CHIKV e EEEV, acessos 6NK7 e 6MX4 no RCSB PDB respectivamente. A análise de qualidade da estrutura da proteína E2 de CHIKV apresentou Clashscore de 16,31 (percentil 97th), MolProbity score de 2,30 (percentil 99th) e Z-score de -5,29. O modelo apresentou ainda 99.9% dos resíduos de aminoácidos em regiões aceitáveis/favoráveis no Ramachandran (Figura 7). A mesma análise para o cristal da proteína E2 de EEEV apontou Clashscore de 7,5 (percentil 97th), MolProbity score de 2,14 (percentil 100th), Z-score de -5,27, sendo 99.5% dos resíduos de aminoácidos em regiões aceitáveis/favoráveis no Ramachandran (Figura 8).

Os valores dos Clashscore e MolProbity score das proteínas E2 de CHIKV e EEEV, ambos em percentis próximos a 100th, demonstram que estas estruturas apresentam boa qualidade, ou seja, o modelo tridimensional apresenta arranjo atômico satisfatório e apto para ser usado nas análises tridimensionais. Segundo estes valores, a estrutura tem relevante probabilidade de corresponder ao modelo natural da estrutura destas glicoproteínas na partícula viral.

Os 40 epítomos preditos estão distribuídos por toda a extensão das proteínas E2, tendo sido representados em todos os domínios (A, B e C) da cadeia peptídica da proteína E2. A ampla distribuição destes epítomos pode ser verificada na proteína E2 do EEEV (Figura 9). Este resultado corrobora a característica altamente imuno-dominante destas proteínas e possibilita também a utilização destes antígenos para o diagnóstico sorológico (BAGNO et al., 2020). Para o CHIKV, embora amplamente distribuídos pela proteína E2, não houve epítomos conservados preditos para células T-citotóxicas (Anexo 8.6).

A partir das estruturas tridimensionais das proteínas E2, os epítomos de Linfócitos B preditadamente expostos foram submetidos ao cálculo de exposição ao solvente no software Pymol. A área exposta ao solvente variou entre 405,32 e 891,68 para os epítomos de CHIKV e entre 434,39 e 809,40 entre os epítomos de EEEV (Anexo 8.8).

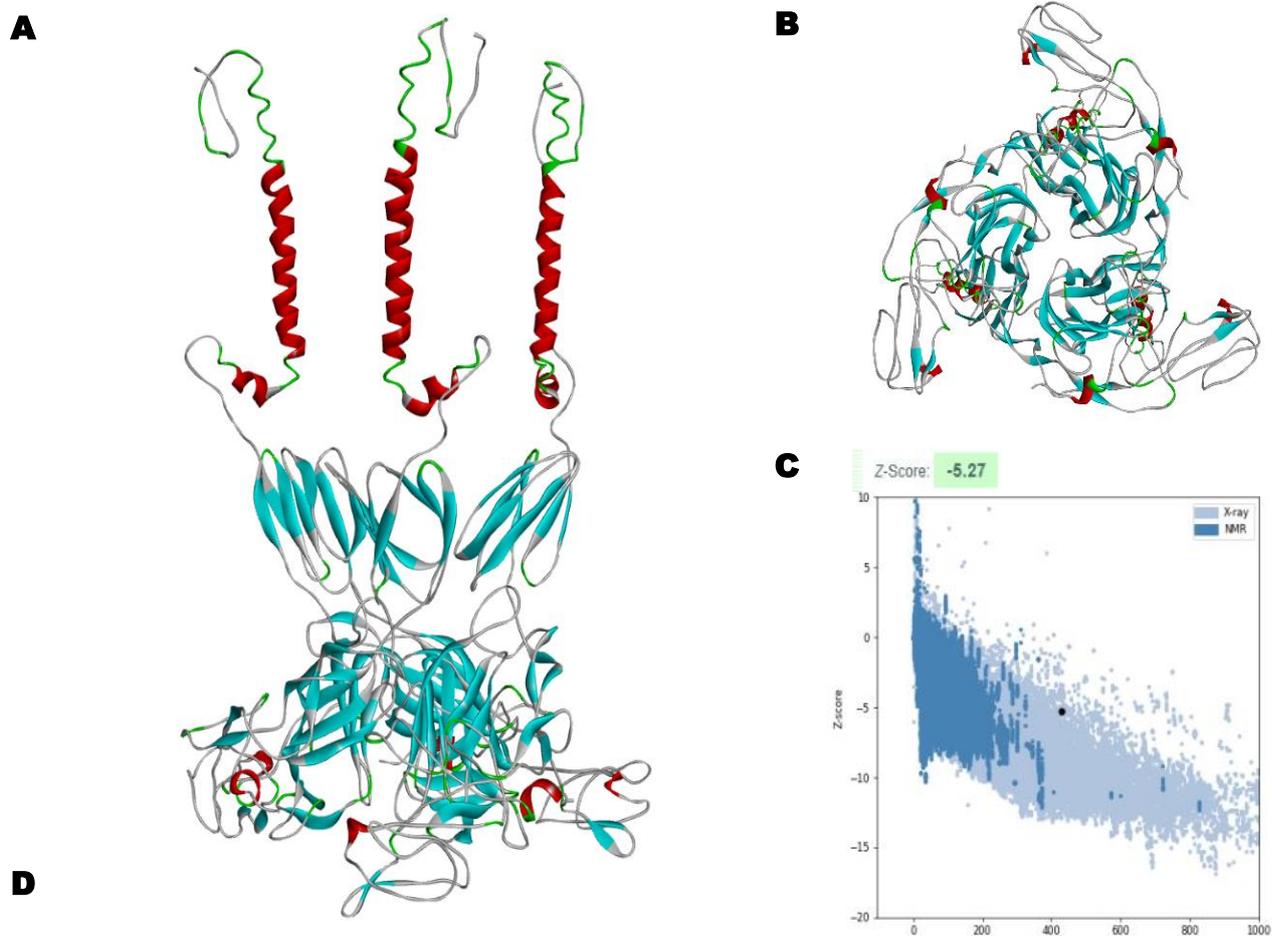
Dentre os 9 epítomos de linfócitos B conservados entre os Alphavirus, apenas o epítopo PGATVPFLSLICCIR de CHIKV foi considerado de topologia interna à membrana pela análise do POA2 e foi retirado das análises estruturais e de cálculo de acessibilidade.



Summary statistics

All-Atom Contacts	Clashscore, all atoms:	16.31	97 th percentile* (N=37, 3Å - 9999Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	32	0.82% Goal: <0.3%
	Favored rotamers	3416	87.91% Goal: >98%
	Ramachandran outliers	4	0.09% Goal: <0.05%
	Ramachandran favored	3980	88.39% Goal: >98%
	Rama distribution Z-score	-4.41 ± 0.10	Goal: abs(Z score) < 2
	MolProbity score ^o	2.30	99 th percentile* (N=342, 3.25Å - 5.24Å)
	Bad bonds:	2 / 36092	0.01% Goal: 0%
Bad angles:	30 / 49135	0.06% Goal: <0.1%	
Additional validations	Chiral volume outliers	0/5499	
	Waters with clashes	0/0	0.00% See UnDowser table for details

Figura 7: Resultados da validação da estrutura tridimensional da proteína E2 de CHIKV (Acesso RCSB PDB: 6NK7 (BASORE et al., 2019)). A) Vista lateral da proteína E2 de CHIKV. B) Vista em rotação de 90° sob o domínio A da proteína E2 de CHIKV. C) e D) Gráfico e tabela resultantes das análises feitas pelo MolProbity e ProSA, respectivamente.



Summary statistics

All-Atom Contacts	Clashscore, all atoms:	7.5	97 th percentile* (N=37, 3Å - 9999Å)
	Clashscore is the number of serious steric overlaps (> 0.4 Å) per 1000 atoms.		
Protein Geometry	Poor rotamers	2	0.06% Goal: <0.3%
	Favored rotamers	3346	97.07% Goal: >98%
	Ramachandran outliers	21	0.52% Goal: <0.05%
	Ramachandran favored	3214	80.21% Goal: >98%
	Rama distribution Z-score	-5.32 ± 0.11	Goal: abs(Z score) < 2
	MolProbity score [^]	2.14	100 th percentile* (N=342, 3.25Å - 4.65Å)
	Bad bonds:	0 / 32167	0.00% Goal: 0%
	Bad angles:	30 / 43840	0.07% Goal: <0.1%
Additional validations	Chiral volume outliers	0/4939	
	Waters with clashes	0/0	0.00% See UnDowser table for details

Figura 8: Resultados da validação da estrutura tridimensional da proteína E2 de EEEV (Acesso RCSB PDB: 6MX4 (HASAN et al., 2018)). A) Vista lateral da proteína E2 de CHIKV. B) Vista em rotação de 90° sob o domínio A da proteína E2 de CHIKV. C) e D) Gráfico e tabela resultantes das análises feitas pelo MolProbity e PROSA, respectivamente.

DISTRIBUIÇÃO DOS EPÍTOPOS DE CÉLULAS B E CÉLULAS T NA PROTEÍNA E2 DE EEV

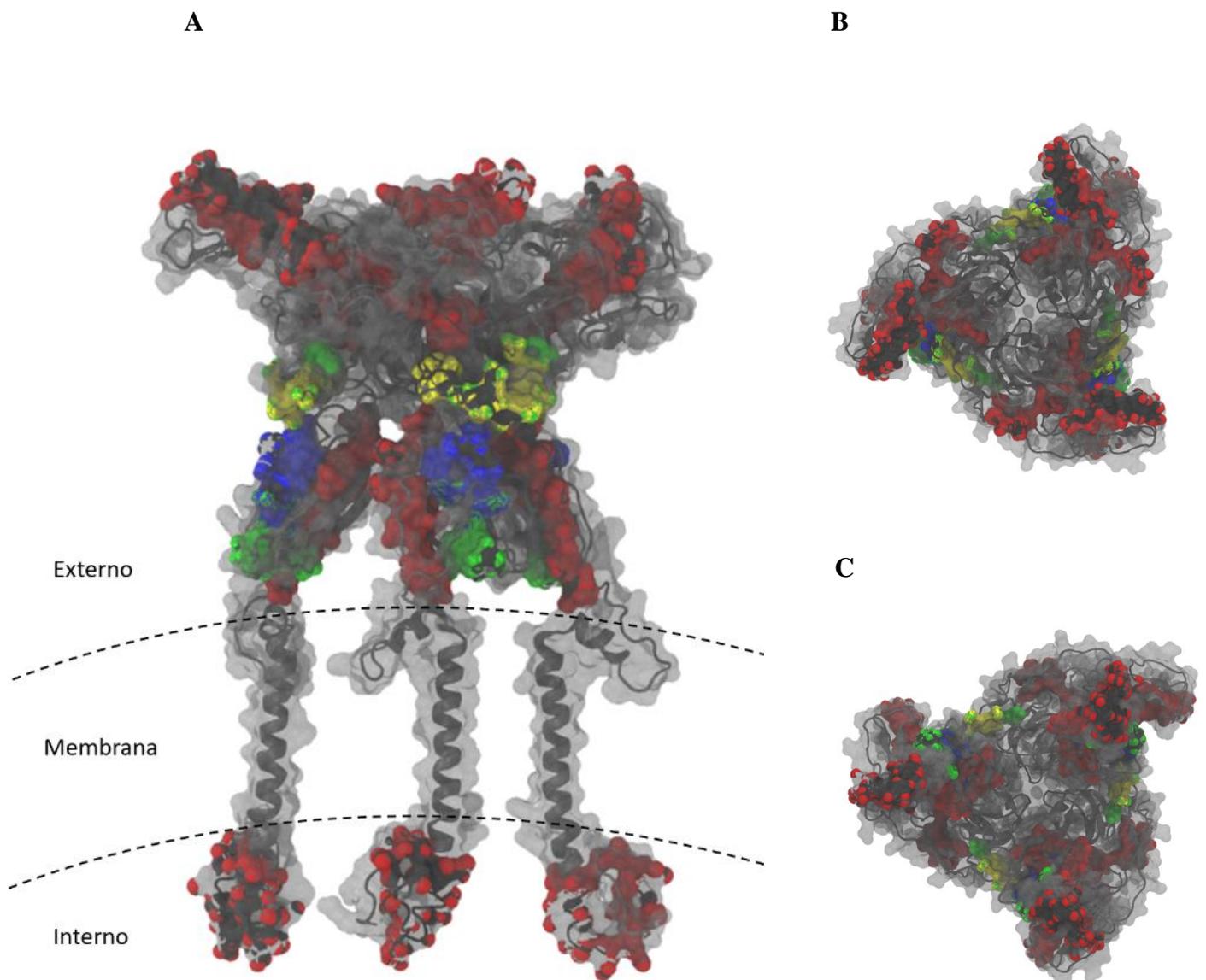


Figura 9: Visualização 3D da distribuição dos epítomos preditos na proteína E2 de EEV, com base no cristal selecionado no banco de dados RCSB PDB (Acesso: 6MX4). Estão caracterizados na estrutura os epítomos lineares preditos para células B pelas ferramentas Bepipred (Azul) e IMED (Amarelo), epítomos preditos para células T-citotóxicas pela ferramenta NetCTL (Verde) e os epítomos de células T-auxiliares, preditos pela ferramenta MHCII Binding Predictions (Vermelho). (A) Vista longitudinal da proteína E2 de EEV. (B) Vista da proteína E2 rotacionada em 90° para vista superior ao domínio A e B. (C) Vista rotacionada em 90° da proteína E2 sob domínio C.

O cálculo de exposição dos epítomos para Linfócitos B foi realizado para 8 dos 9 epítomos preditos para os vírus CHIKV e EEEV. A predição de exposição nas proteínas virais contribui para apontar os alvos realmente acessíveis aos componentes do sistema imune. A escolha por epítomos expostos busca mimetizar o estado natural da cápsula viral, onde a superfície expõe proteínas e/ou peptídeos imunodominantes capazes de elicitar respostas de anticorpos específicos e também diferentes respostas de células T (RAHMAN et al., 2020).

A predição de epítomos para mais de um tipo celular da resposta imune permite a construção de plataformas vacinais mais eficientes. Neste estudo, selecionamos 39 epítomos lineares, conservados entre a proteína estrutural E2 dos vírus CHIKV e EEEV, isentos de potencial tóxico ou alergênico e com relevante potencial para se tornarem alvos para protótipos vacinais. Estes epítomos foram selecionados por diferentes algoritmos de predição de epítomos e triados por duas ferramentas computacionais, POA1 e POA2.

O desenho de múltiplos-epítomos requer demais predições que não estão incluídas no presente estudo, como por exemplo, análise de cobertura populacional, docking molecular, caracterização físico-química, entre outras. Todavia, a construção de vacinas quiméricas é ainda promissora, biologicamente segura e economicamente atrativa pois não necessita de culturas in vitro de vírus, possui menor tempo de desenvolvimento e garantem uma ativação eficaz da resposta imune (L. DUDEK et al., 2010; RAHMAN et al., 2020; WANG et al., 2019).

O uso futuro dos alvos apontados em associação com linkers e adjuvantes podem aumentar a sua imunogenicidade através da indução da resposta inata do sistema imune, e por consequência, ativar a resposta imune adaptativa, que são fundamentais para o design de abordagens vacinais promissoras (SANCHEZ-TRINCADO; GOMEZ-PEROSANZ; RECHE, 2017).

O resultado apresentado pelo POA1 e POA2 correspondem a uma parte fundamental das análises necessárias para construção de uma quimera de múltiplos-epítomos, a predição e seleção de bons alvos antigênicos. Desta forma, o presente trabalho viabiliza o uso de uma plataforma computacional multi-integrada para a otimização da produção vacinais e testes diagnósticos. Oferecendo protótipos putativos de forma eficiente, rápida e menos laboriosa. O POA permite ser melhorado e ampliado, com a inclusão de outros módulos de predição e análise

que permitam incluir mais etapas e/ou critérios à seleção dos epítomos para protótipos vacinais e de testes diagnósticos.

6. CONSIDERAÇÕES FINAIS

Neste trabalho selecionamos epítomos conservados entre as proteínas E2 dos vírus CHIKV e EEEV visando a construção de um protótipo vacinal, através do uso das ferramentas POA, um pipeline semi-automático de Otimização de Antígenos, com o intuito de facilitar a etapa de seleção de alvos imunogênicos no desenvolvimento de abordagens vacinais e testes diagnósticos.

Aplicando este pipeline à prospecção de alvos em *Alphavirus* de interesse médico, foram encontrados 8 epítomos de células B, 2 de células T-citotóxicas e 29 de células T-auxiliares lineares e conservados. Estes epítomos poderão ser utilizados para o desenvolvimento de abordagem vacinal baseada em múltiplos-epítomos.

Como pesquisa pioneira no desenvolvimento de ferramentas computacionais dentro do nosso grupo de pesquisa, este trabalho abre espaço para o desenvolvimento de novos trabalhos de imunoinformática com patógenos virais e não-virais. Além disso, novas versões podem incluir outras etapas de predição no pipeline e permitir a utilização destes dados em outras abordagens *in silico* no desenvolvimento de vacinas e testes diagnósticos.

Os planos de melhoria em versões futuras do POA incluem a exclusão da etapa manual de predição de conservação, a inclusão de módulos para outros algoritmos de predição de antigenicidade e o desenvolvimento de módulos para outras análises, como predições de toxicidade, alergenicidade e glicosilações. Tais melhorias podem facilitar ainda mais a escolha de bons alvos imunogênicos e otimizar o desenho de quimeras proteicas visando abordagens vacinais e testes diagnósticos, favorecendo os avanços no diagnóstico e controle de doenças virais e não-virais.

7. REFERÊNCIAS BIBLIOGRÁFICAS

ALMOFTI, Yassir A.; ABD-ELRAHMAN, Khoubieb Ali; ELTILIB, Elsideeq E. M. Vaccinomic approach for novel multi epitopes vaccine against severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). **BMC Immunology**, [S. l.], v. 22, n. 1, p. 1–20, 2021. DOI: 10.1186/S12865-021-00412-0/TABLES/4. Disponível em: <https://bmcimmunol.biomedcentral.com/articles/10.1186/s12865-021-00412-0>. Acesso em: 8 jun. 2022.

ARBOVIROSES. Direção de Rafael Figueiredo. Produção de Christovão Paiva. Roteiro: Marcela Morato. Rio de Janeiro: Canal Saúde Fiocruz, 2017. 1 vídeo MPEG-4 (26 min.), son., color. Série Ligado em Saúde. Disponível em: <https://www.canalsaude.fiocruz.br/canal/videoAberto/arboviroses-LES-1924>. Acesso em: 13 maio 2022.

BAE, Kyungdong; CHOI, Junyoul; JANG, Yangsuk; AHN, Sangjeom; HUR, Byungki. Innovative vaccine production technologies: The evolution and value of vaccine production technologies. **Archives of Pharmacal Research** 2009 32:4, [S. l.], v. 32, n. 4, p. 465–480, 2009. DOI: 10.1007/S12272-009-1400-1. Disponível em: <https://link.springer.com/article/10.1007/s12272-009-1400-1>. Acesso em: 9 jun. 2022.

BAGNO, Flávia Fonseca; GODÓI, Lara Carvalho; FIGUEIREDO, Maria Marta; RODRIGUES SÉRGIO, Sarah Aparecida; DE FÁTIMA SILVA MORAES, Thaís; DE CASTRO SALAZAR, Natália; KIM, Young Chan; REYES-SANDOVAL, Arturo; GUIMARÃES DA FONSECA, Flávio. Chikungunya E2 Protein Produced in E. coli and HEK293-T Cells—Comparison of Their Performances in ELISA. **Viruses** 2020, Vol. 12, Page 939, [S. l.], v. 12, n. 9, p. 939, 2020. DOI: 10.3390/V12090939. Disponível em: <https://www.mdpi.com/1999-4915/12/9/939/htm>. Acesso em: 8 jun. 2022.

BASORE, Katherine et al. Cryo-EM Structure of Chikungunya Virus in Complex with the Mxra8 Receptor. **Cell**, [S. l.], v. 177, n. 7, p. 1725- 1737.e16, 2019. DOI: 10.1016/J.CELL.2019.04.006/ATTACHMENT/F4025194-0F8F-4AF0-AC71-60778E46C64C/MMC1.PDF. Disponível em:

<http://www.cell.com/article/S0092867419303927/fulltext>. Acesso em: 25 maio. 2022.

BUI, Huynh Hoa; SIDNEY, John; LI, Wei; FUSSEDER, Nicolas; SETTE, Alessandro. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. **BMC bioinformatics**, [S. l.], v. 8, 2007. DOI: 10.1186/1471-2105-8-361. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17897458/>. Acesso em: 7 mar. 2022.

BYRD, Emily A.; KIELIAN, Margaret. An Alphavirus E2 Membrane-Proximal Domain Promotes Envelope Protein Lateral Interactions and Virus Budding. **mBio**, [S. l.], v. 8, n. 6, 2017. DOI: 10.1128/MBIO.01564-17. Disponível em: </pmc/articles/PMC5676042/>. Acesso em: 9 jun. 2022.

CAICEDO, Edgar Yaset; CHARNIGA, Kelly; RUEDA, Amanecer; DORIGATTI, Ilaria; MENDEZ, Yardany; HAMLET, Arran; CARRERA, Jean Paul; CUCUNUBÁ, Zulma M. The epidemiology of Mayaro virus in the Americas: A systematic review and key parameter estimates for outbreak modelling. **PLOS Neglected Tropical Diseases**, [S. l.], v. 15, n. 6, p. e0009418, 2021. DOI: 10.1371/JOURNAL.PNTD.0009418. Disponível em: <https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0009418>. Acesso em: 6 jan. 2022.

CARRERA, Jean-Paul et al. Eastern Equine Encephalitis in Latin America. **New England Journal of Medicine**, [S. l.], v. 369, n. 8, p. 732–744, 2013. DOI: 10.1056/NEJMOA1212628/SUPPL_FILE/NEJMOA1212628_DISCLOSURES.PDF. Disponível em: <https://www.nejm.org/doi/10.1056/NEJMoa1212628>. Acesso em: 18 maio. 2022.

CHEN, Rubing et al. ICTV virus taxonomy profile: Togaviridae. **Journal of General Virology**, [S. l.], v. 99, n. 6, p. 761–762, 2018. DOI: 10.1099/JGV.0.001072/CITE/REFWORKS. Disponível em: <https://www.microbiologyresearch.org/content/journal/jgv/10.1099/jgv.0.001072>. Acesso em: 20 maio. 2022.

CONWAY, Michael J.; COLPITTS, Tonya M.; FIKRIG, Erol. Role of the Vector in

Arbovirus Transmission. **Annual review of virology**, [S. l.], v. 1, n. 1, p. 71, 2014. DOI: 10.1146/ANNUREV-VIROLOGY-031413-085513. Disponível em: /pmc/articles/PMC7809425/. Acesso em: 6 jan. 2022.

DEL TORDELLO, E.; RAPPUOLI, R.; DELANY, I. Reverse Vaccinology: Exploiting Genomes for Vaccine Design. In: MODJARRAD, K.; KOFF, W. C. (org.). **Human Vaccines: Emerging Technologies in Design and Development**. [s.l.] : Elsevier Inc., 2017. p. 65–86. DOI: 10.1016/B978-0-12-802302-0.00002-9.

DIAGNE, Cheikh Tidiane; BENGUE, Michèle; CHOUMET, Valérie; HAMEL, Rodolphe; POMPON, Julien; MISSÉ, Dorothée. Mayaro Virus Pathogenesis and Transmission Mechanisms. **Pathogens** 2020, Vol. 9, Page 738, [S. l.], v. 9, n. 9, p. 738, 2020. DOI: 10.3390/pathogens9090738. Disponível em: www.mdpi.com/journal/pathogens.

FEASEY, Nick; WANSBROUGH-JONES, Mark; MABEY, David C. W.; SOLOMON, Anthony W. Neglected tropical diseases. **British Medical Bulletin**, [S. l.], v. 93, n. 1, p. 179–200, 2010. DOI: 10.1093/BMB/LDP046. Disponível em: https://academic.oup.com/bmb/article/93/1/179/307584. Acesso em: 10 maio. 2022.

FISCH, Gene S. Visual inspection of data revisited: Do the eyes still have it? **The Behavior Analyst**, [S. l.], v. 21, n. 1, p. 111, 1998. DOI: 10.1007/BF03392786. Disponível em: /pmc/articles/PMC2731381/?report=abstract. Acesso em: 9 mar. 2022.

FLOREA, L.; HALLDÓRSSON, B.; KOHLBACHER, O.; SCHWARTZ, R.; HOFFMAN, S.; ISTRAIL, S. Epitope prediction algorithms for peptide-based vaccine design. **Proceedings of the 2003 IEEE Bioinformatics Conference, CSB 2003**, [S. l.], p. 17–26, 2003. DOI: 10.1109/CSB.2003.1227293.

FURUYA-KANAMORI, Luis et al. Co-distribution and co-infection of chikungunya and dengue viruses. **BMC Infectious Diseases**, [S. l.], v. 16, n. 1, p. 1–11, 2016. DOI: 10.1186/S12879-016-1417-2/FIGURES/2. Disponível em: https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-016-1417-2. Acesso em: 9 jun. 2022.

GAO, Shan; SONG, Siqu; ZHANG, Leiliang. Recent Progress in Vaccine Development Against Chikungunya Virus. **Frontiers in Microbiology**, [S. l.], v. 10, p. 2881, 2019. DOI: 10.3389/FMICB.2019.02881/BIBTEX.

GONÇALVES, Ricardo Lemes et al. **SARS-CoV-2 mutations and where to find them: an in silico perspective of structural changes and antigenicity of the spike protein** **Journal of Biomolecular Structure and Dynamics** Taylor and Francis Ltd., , 2020. DOI: 10.1080/07391102.2020.1844052. Disponível em: <https://www.tandfonline.com/doi/full/10.1080/07391102.2020.1844052>. Acesso em: 4 mar. 2021.

GOULD, E. A.; HIGGS, S. Impact of climate change and other factors on emerging arbovirus diseases. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, [S. l.], v. 103, n. 2, p. 109–121, 2009. DOI: 10.1016/j.trstmh.2008.07.025.

GUPTA R.; JUNG E.; BRUNAK S. Prediction of N-glycosylation sites in human proteins. **In preparation**, [S. l.], 2004. Disponível em: <https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0>. Acesso em: 4 mar. 2021.

GUPTA, Sudheer; KAPOOR, Pallavi; CHAUDHARY, Kumardeep; GAUTAM, Ankur; KUMAR, Rahul; RAGHAVA, Gajendra P. S. In Silico Approach for Predicting Toxicity of Peptides and Proteins. **PLOS ONE**, [S. l.], v. 8, n. 9, p. e73957, 2013. DOI: 10.1371/JOURNAL.PONE.0073957. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0073957>. Acesso em: 7 jan. 2022.

HASAN, S. Saif et al. Cryo-EM Structures of Eastern Equine Encephalitis Virus Reveal Mechanisms of Virus Disassembly and Antibody Neutralization. **Cell Reports**, [S. l.], v. 25, n. 11, p. 3136-3147.e5, 2018. DOI: 10.1016/J.CELREP.2018.11.067/ATTACHMENT/F8526F0D-B96F-46E5-A54C-2C088B1A7A75/MMC1.PDF. Disponível em: <http://www.cell.com/article/S2211124718318436/fulltext>. Acesso em: 25 maio. 2022.

JESPERSEN, Martin Closter; PETERS, Bjoern; NIELSEN, Morten; MARCATILI, Paolo. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. **Nucleic acids research**, [S. l.], v. 45, n. W1, p. W24–W29, 2017. DOI: 10.1093/NAR/GKX346. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/28472356/>. Acesso em: 7 jan. 2022.

JIANG, Pengfei; CAI, Yiqi; CHEN, Jun; YE, Xiaoxian; MAO, Shanshan; ZHU, Shanli; XUE, Xiangyang; CHEN, Shao; ZHANG, Lifang. Evaluation of tandem Chlamydia trachomatis MOMP multi-epitopes vaccine in BALB/c mice model. **Vaccine**, [S. l.], v. 35, n. 23, p. 3096–3103, 2017. DOI: 10.1016/J.VACCINE.2017.04.031.

KANDUC, Darja. Peptimmunology: Immunogenic Peptides and Sequence Redundancy. **Current Drug Discovery Technologies**, [S. l.], v. 2, n. 4, p. 239–244, 2005. DOI: 10.2174/157016305775202946.

KHAN, Afjal Hossain; MORITA, Kouichi; DEL CARMEN PARQUET, Maria; HASEBE, Futoshi; MATHENGE, Edward G. M.; IGARASHI, Akira. Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. **The Journal of general virology**, [S. l.], v. 83, n. Pt 12, p. 3075–3084, 2002. DOI: 10.1099/0022-1317-83-12-3075. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/12466484/>. Acesso em: 27 jun. 2022.

KIM, Arthur S. et al. Pan-protective anti-alphavirus human antibodies target a conserved E1 protein epitope. **Cell**, [S. l.], v. 184, n. 17, p. 4414–4429.e19, 2021. DOI: 10.1016/J.CELL.2021.07.006.

KOLASKAR, A. S.; TONGAONKAR, Prasad C. A semi-empirical method for prediction of antigenic determinants on protein antigens. **FEBS Letters**, [S. l.], v. 276, n. 1–2, p. 172–174, 1990. DOI: 10.1016/0014-5793(90)80535-Q. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/1702393/>. Acesso em: 4 mar. 2021.

KROGH, Anders; È RN LARSSON, Bjo; VON HEIJNE, Gunnar; L SONNHAMMER,

Erik L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. **Journal of Molecular Biology**, [S. l.], v. 305, n. 3, 2001. DOI: 10.1006/jmbi.2000.4315. Disponível em: <http://www.cbs.dtu.dk/services/TMHMM/>. Acesso em: 4 mar. 2021.

KROON CAMPOS, Rafael; PRECIADO-LLANES, Lorena; AZAR, Sasha R.; KIM, Young Chan; BRANDON, Olivia; LÓPEZ-CAMACHO, César; REYES-SANDOVAL, Arturo; ROSSI, Shannan L. Adenoviral-Vectored Mayaro and Chikungunya Virus Vaccine Candidates Afford Partial Cross-Protection From Lethal Challenge in A129 Mouse Model. **Frontiers in Immunology**, [S. l.], v. 11, p. 2802, 2020. DOI: 10.3389/FIMMU.2020.591885/BIBTEX.

KUMAR, Sudhir; STECHER, Glen; LI, Michael; KNYAZ, Christina; TAMURA, Koichiro. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. **Molecular Biology and Evolution**, [S. l.], v. 35, n. 6, p. 1547, 2018. DOI: 10.1093/MOLBEV/MSY096. Disponível em: </pmc/articles/PMC5967553/>. Acesso em: 29 maio. 2022.

L. DUDEK, Nadine; PERLMUTTER, Patrick; ISABEL AGUILAR, Marie-; P. CROFT, Nathan; W. PURCELL, Anthony. Epitope Discovery and Their Use in Peptide Based Vaccines. **Current Pharmaceutical Design**, [S. l.], v. 16, n. 28, p. 3149–3157, 2010. DOI: 10.2174/138161210793292447.

LANGSJOEN, Rose M. et al. Chikungunya Virus Strains Show Lineage-Specific Variations in Virulence and Cross-Protective Ability in Murine and Nonhuman Primate Models. **mBio**, [S. l.], v. 9, n. 2, 2018. DOI: 10.1128/MBIO.02449-17. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29511072/>. Acesso em: 6 jan. 2022.

LARSEN, Mette V.; LUNDEGAARD, Claus; LAMBERTH, Kasper; BUUS, Soren; LUND, Ole; NIELSEN, Morten. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. **BMC bioinformatics**, [S. l.], v. 8, 2007. DOI: 10.1186/1471-2105-8-424. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17973982/>. Acesso em: 7 jan. 2022.

LEUNG, Jason Yat Sing; NG, Mary Mah Lee; CHU, Justin Jang Hann. Replication of alphaviruses: A review on the entry process of alphaviruses into cells. **Advances in Virology**, [S. l.], v. 2011, 2011. DOI: 10.1155/2011/249640.

LI, Long; JOSE, Joyce; XIANG, Ye; KUHN, Richard J.; ROSSMANN, Michael G. Structural changes of envelope proteins during alphavirus fusion. **Nature** 2010 **468:7324**, [S. l.], v. 468, n. 7324, p. 705–708, 2010. DOI: 10.1038/nature09546. Disponível em: <https://www.nature.com/articles/nature09546>. Acesso em: 8 jun. 2022.

LI, Shi Hua; LI, Xiao Feng; ZHAO, Hui; JIANG, Tao; DENG, Yong Qiang; YU, Xue Dong; ZHU, Qing Yu; QIN, E. De; QIN, Cheng Feng. Cross protection against lethal West Nile virus challenge in mice immunized with recombinant E protein domain III of Japanese encephalitis virus. **Immunology Letters**, [S. l.], v. 138, n. 2, p. 156–160, 2011. DOI: 10.1016/J.IMLET.2011.04.003.

LINDOSO, J. A. L.; LINDOSO, A. A. B. Doenças tropicais negligenciadas no Brasil . **Revista do Instituto de Medicina Tropical de São Paulo**, [S. l.], v. 51, n. 5, p. 247-253, 2009. Disponível em: <https://www.revistas.usp.br/rimtsp/article/view/31284>. Acesso em: 10 maio. 2022.

LUNDSTROM, Kenneth. Alphavirus-Based Vaccines. **Viruses** 2014, Vol. 6, Pages **2392-2415**, [S. l.], v. 6, n. 6, p. 2392–2415, 2014. DOI: 10.3390/V6062392. Disponível em: <https://www.mdpi.com/1999-4915/6/6/2392/htm>. Acesso em: 6 jan. 2022.

MAURER-STROH, Sebastian; KRUTZ, Nora L.; KERN, Petra S.; GUNALAN, Vithiagarar; NGUYEN, Minh N.; LIMVIPHUVADH, Vachiranee; EISENHABER, Frank; FRANK GERBERICK, G. AllerCatPro-prediction of protein allergenicity potential from the protein sequence. **Bioinformatics (Oxford, England)**, [S. l.], v. 35, n. 17, p. 3020–3027, 2019. DOI: 10.1093/BIOINFORMATICS/BTZ029. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/30657872/>. Acesso em: 7 jan. 2022.

MEIJERING, Erik; DZYUBACHYK, Oleh; SMAL, Ihor; VAN CAPPELLEN, Wiggert A. Tracking in cell and developmental biology. **Seminars in Cell & Developmental Biology**,

[S. l.], v. 20, n. 8, p. 894–902, 2009. DOI: 10.1016/J.SEMCDB.2009.07.004.

MENDES, Adriano; KUHN, Richard J. Alphavirus Nucleocapsid Packaging and Assembly. **Viruses** 2018, Vol. 10, Page 138, [S. l.], v. 10, n. 3, p. 138, 2018. DOI: 10.3390/V10030138. Disponível em: <https://www.mdpi.com/1999-4915/10/3/138/htm>. Acesso em: 6 jan. 2022.

MODJARRAD, Kayvon; KOFF, Wayne C. (ORG.). **Human vaccines : emerging technologies in design and development**. 1º Edição ed. [s.l.] : Academic Press, 2016.

MOLYNEUX, David. Neglected tropical diseases. **Community Eye Health**, [S. l.], v. 26, n. 82, p. 21, 2013. Disponível em: </pmc/articles/PMC3756642/>. Acesso em: 10 maio. 2022.

NGUYEN, Wilson et al. Arthritogenic Alphavirus Vaccines: Serogrouping Versus Cross-Protection in Mouse Models. **Vaccines** 2020, Vol. 8, Page 209, [S. l.], v. 8, n. 2, p. 209, 2020. DOI: 10.3390/VACCINES8020209. Disponível em: <https://www.mdpi.com/2076-393X/8/2/209/htm>. Acesso em: 9 jun. 2022.

OLIVIA, Lwande Wesula; OBANDA, Vincent; BUCHT, Göran; MOSOMTAI, Gladys; OTIENO, Viola; AHLM, Clas; EVANDER, Magnus. Global emergence of Alphaviruses that cause arthritis in humans. <https://doi.org/10.3402/iee.v5.29853>, [S. l.], v. 5, n. 1, p. 29853, 2015. DOI: 10.3402/IEE.V5.29853. Disponível em: <https://www.tandfonline.com/doi/abs/10.3402/iee.v5.29853>. Acesso em: 6 jan. 2022.

PETERS, Bjoern; NIELSEN, Morten; SETTE, Alessandro. T Cell Epitope Predictions. **Annual Review of Immunology**, [S. l.], v. 38, p. 123–145, 2020. DOI: 10.1146/ANNUREV-IMMUNOL-082119-124838.

PINHEIRO, Gilson Marques. **VARIABILIDADES DOS DEFEITOS DO PRODUTO E DESEMPENHO DO INSPECIONISTA**. 2007. [S. l.], 2007.

RAHMAN, M. Shaminur et al. Epitope-based chimeric peptide vaccine design against S, M and e proteins of SARS-CoV-2, the etiologic agent of COVID-19 pandemic: An in silico

approach. **PeerJ**, [S. l.], v. 8, p. e9572, 2020. DOI: 10.7717/PEERJ.9572/SUPP-16. Disponível em: <https://peerj.com/articles/9572>. Acesso em: 10 abr. 2022.

RANDOLPH, Sarah E.; ROGERS, David J. The arrival, establishment and spread of exotic diseases: patterns and predictions. **Nature Reviews Microbiology** 2010 **8:5**, [S. l.], v. 8, n. 5, p. 361–371, 2010. DOI: 10.1038/nrmicro2336. Disponível em: <https://www.nature.com/articles/nrmicro2336>. Acesso em: 10 maio. 2022.

RANGEL, Margarita V; STAPLEFORD, Kenneth A.; MUKHOPADHYAY, Tuli; MORRISON, Tem. Alphavirus Virulence Determinants. **Pathogens** 2021, Vol. 10, Page 981, [S. l.], v. 10, n. 8, p. 981, 2021. DOI: 10.3390/PATHOGENS10080981. Disponível em: <https://www.mdpi.com/2076-0817/10/8/981/htm>. Acesso em: 7 maio. 2022.

RAO, Shambhavi; TAYLOR, Adam. Arthritogenic Alphavirus Capsid Protein. **Life** 2021, Vol. 11, Page 230, [S. l.], v. 11, n. 3, p. 230, 2021. DOI: 10.3390/LIFE11030230. Disponível em: <https://www.mdpi.com/2075-1729/11/3/230/htm>. Acesso em: 6 jan. 2022.

RAPPUOLI, Rino. Reverse vaccinology. **Current Opinion in Microbiology**, [S. l.], v. 3, n. 5, p. 445–450, 2000. DOI: 10.1016/S1369-5274(00)00119-3.

RAPPUOLI, Rino; BOTTOMLEY, Matthew J.; D'ORO, Ugo; FINCO, Oretta; DE GREGORIO, Ennio. Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. **Journal of Experimental Medicine**, [S. l.], v. 213, n. 4, p. 469–481, 2016. DOI: 10.1084/JEM.20151960. Disponível em: www.jem.org/cgi/doi/10.1084/jem.20151960469. Acesso em: 13 maio. 2022.

RYMAN, Kate D.; KLIMSTRA, William B. Host responses to alphavirus infection. **Immunological Reviews**, [S. l.], v. 225, n. 1, p. 27–45, 2008. DOI: 10.1111/J.1600-065X.2008.00670.X. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1600-065X.2008.00670.x>. Acesso em: 6 jan. 2022.

RYU, Wang-Shick. Virus Structure. **Molecular Virology of Human Pathogenic Viruses**, [S. l.], p. 21–29, 2017. DOI: 10.1016/B978-0-12-800838-6.00002-3. Disponível em:

<https://linkinghub.elsevier.com/retrieve/pii/B9780128008386000023>. Acesso em: 10 maio. 2022.

SANCHEZ-TRINCADO, Jose L.; GOMEZ-PEROSANZ, Marta; RECHE, Pedro A. Fundamentals and Methods for T- and B-Cell Epitope Prediction. **Journal of Immunology Research**, [S. l.], v. 2017, 2017. DOI: 10.1155/2017/2680160.

SEMENZA, Jan C.; SUK, Jonathan E. Vector-borne diseases and climate change: a European perspective. **FEMS microbiology letters**, [S. l.], v. 365, n. 2, 2018. DOI: 10.1093/FEMSLE/FNX244. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29149298/>. Acesso em: 10 maio. 2022.

SETTE, Alessandro; RAPPUOLI, Rino. Reverse Vaccinology: Developing Vaccines in the Era of Genomics. **Immunity**, [S. l.], v. 33, n. 4, p. 530, 2010. DOI: 10.1016/J.IMMUNI.2010.09.017. Disponível em: </pmc/articles/PMC3320742/>. Acesso em: 9 jun. 2022.

SHARMA, Anuj; KNOLLMANN-RITSCHER, Barbara. Current Understanding of the Molecular Basis of Venezuelan Equine Encephalitis Virus Pathogenesis and Vaccine Development. **Viruses 2019, Vol. 11, Page 164**, [S. l.], v. 11, n. 2, p. 164, 2019. DOI: 10.3390/V11020164. Disponível em: <https://www.mdpi.com/1999-4915/11/2/164/htm>. Acesso em: 6 jan. 2022.

SHEY, Robert Adamu et al. In-silico design of a multi-epitope vaccine candidate against onchocerciasis and related filarial diseases. **Scientific Reports 2019 9:1**, [S. l.], v. 9, n. 1, p. 1–18, 2019. DOI: 10.1038/s41598-019-40833-x. Disponível em: <https://www.nature.com/articles/s41598-019-40833-x>. Acesso em: 8 jun. 2022.

SONNHAMMER, E. L.; VON HEIJNE, G.; KROGH, A. A hidden Markov model for predicting transmembrane helices in protein sequences. **Proceedings. International Conference on Intelligent Systems for Molecular Biology**, [S. l.], v. 6, p. 175–182, 1998. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/9783223/>. Acesso em: 25 maio. 2022.

SORIA-GUERRA, Ruth E.; NIETO-GOMEZ, Ricardo; GOVEA-ALONSO, Dania O.; ROSALES-MENDOZA, Sergio. An overview of bioinformatics tools for epitope prediction: Implications on vaccine development. **Journal of Biomedical Informatics**, [S. l.], v. 53, p. 405–414, 2015. DOI: 10.1016/J.JBI.2014.11.003.

STEENTOFT, Catharina et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. **EMBO Journal**, [S. l.], v. 32, n. 10, p. 1478–1488, 2013. DOI: 10.1038/emboj.2013.79. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/23584533/>. Acesso em: 4 mar. 2021.

STOTHARD, P. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. **BioTechniques**, [S. l.], v. 28, n. 6, 2000. DOI: 10.2144/00286IR01. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/10868275/>. Acesso em: 27 jun. 2022.

STROMBERG, Zachary R.; FISCHER, Will; BRADFUTE, Steven B.; KUBICEK-SUTHERLAND, Jessica Z.; HRABER, Peter. Vaccine Advances against Venezuelan, Eastern, and Western Equine Encephalitis Viruses. **Vaccines 2020, Vol. 8, Page 273**, [S. l.], v. 8, n. 2, p. 273, 2020. DOI: 10.3390/VACCINES8020273. Disponível em: <https://www.mdpi.com/2076-393X/8/2/273/htm>. Acesso em: 6 jan. 2022.

SUHRBIER, Andreas; JAFFAR-BANDJEE, Marie Christine; GASQUE, Philippe. Arthritogenic alphaviruses—an overview. **Nature Reviews Rheumatology 2012 8:7**, [S. l.], v. 8, n. 7, p. 420–429, 2012. DOI: 10.1038/nrrheum.2012.64. Disponível em: <https://www.nature.com/articles/nrrheum.2012.64>. Acesso em: 7 mar. 2022.

THOMPSON, Julie D.; HIGGINS, Desmond G.; GIBSON, Toby J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic acids research**, [S. l.], v. 22, n. 22, p. 4673–4680, 1994. DOI: 10.1093/NAR/22.22.4673. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/7984417/>. Acesso em: 29 maio. 2022.

URA, Takehiro; YAMASHITA, Akio; MIZUKI, Nobuhisa; OKUDA, Kenji;

SHIMADA, Masaru. New vaccine production platforms used in developing SARS-CoV-2 vaccine candidates. **Vaccine**, [S. l.], v. 39, n. 2, p. 197–201, 2021. DOI: 10.1016/J.VACCINE.2020.11.054.

VANEY, Marie Christine; DUQUERROY, Stéphane; REY, Félix A. Alphavirus structure: activation for entry at the target cell surface. **Current Opinion in Virology**, [S. l.], v. 3, n. 2, p. 151–158, 2013. DOI: 10.1016/J.COVIRO.2013.04.003.

VOLCHKOV, V. E.; VOLCHKOVA, V. A.; NETESOV, S. V. [Complete nucleotide sequence of the Eastern equine encephalomyelitis virus genome]. **Molekuliarnaia Genetika, Mikrobiologiia i Virusologiia**, [S. l.], n. 5, p. 8–15, 1991. Disponível em: <https://europepmc.org/article/med/1896061>. Acesso em: 27 jun. 2022.

VOSS, James E.; VANEY, Marie Christine; DUQUERROY, Stéphane; VONRHEIN, Clemens; GIRARD-BLANC, Christine; CRUBLET, Elodie; THOMPSON, Andrew; BRICOGNE, Gérard; REY, Félix A. Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography. **Nature** 2010 **468:7324**, [S. l.], v. 468, n. 7324, p. 709–712, 2010. DOI: 10.1038/nature09555. Disponível em: <https://www.nature.com/articles/nature09555>. Acesso em: 8 jun. 2022.

WEAVER, Scott C.; REISEN, William K. Present and future arboviral threats. **Antiviral Research**, [S. l.], v. 85, n. 2, p. 328–345, 2010. DOI: 10.1016/J.ANTIVIRAL.2009.10.008.

WEISS, Christopher M.; LIU, Hongwei; RIEMERSMA, Kasen K.; BALL, Erin E.; COFFEY, Lark L. Engineering a fidelity-variant live-attenuated vaccine for chikungunya virus. **npj Vaccines** 2020 **5:1**, [S. l.], v. 5, n. 1, p. 1–13, 2020. DOI: 10.1038/s41541-020-00241-z. Disponível em: <https://www.nature.com/articles/s41541-020-00241-z>. Acesso em: 6 jan. 2022.

WHO. **Immunization Agenda 2030: A Global Strategy To Leave No One Behind**. 2020a. Disponível em: <https://www.who.int/publications/m/item/immunization-agenda-2030-a-global-strategy-to-leave-no-one-behind>. Acesso em: 8 jun. 2022.

WHO. **Chikungunya fact sheet**. 2020b. Disponível em: <https://www.who.int/news->

room/fact-sheets/detail/chikungunya. Acesso em: 27 jun. 2022.

WHO. **Vaccines and immunization: What is vaccination?** 2021. Disponível em: <https://www.who.int/news-room/questions-and-answers/item/vaccines-and-immunization-what-is-vaccination>. Acesso em: 8 jun. 2022.

WIEDERSTEIN, Markus; SIPPL, Manfred J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. **Nucleic Acids Research**, [S. l.], v. 35, n. suppl_2, p. W407–W410, 2007. DOI: 10.1093/NAR/GKM290. Disponível em: https://academic.oup.com/nar/article/35/suppl_2/W407/2920938. Acesso em: 7 jan. 2022.

WILLIAMS, Christopher J. et al. MolProbity: More and better reference data for improved all-atom structure validation. **Protein science : a publication of the Protein Society**, [S. l.], v. 27, n. 1, p. 293–315, 2018. DOI: 10.1002/PRO.3330. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29067766/>. Acesso em: 7 jan. 2022.

YANG, Xingdong; YU, Xinglong. An introduction to epitope prediction methods and software. **Rev. Med. Virol**, [S. l.], v. 19, p. 77–96, 2009. DOI: 10.1002/rmv.602. Disponível em: www.interscience.wiley.com.

YANG, Zikun; BOGDAN, Paul; NAZARIAN, Shahin. An in silico deep learning approach to multi-epitope vaccine design: a SARS-CoV-2 case study. **Scientific Reports 2021 11:1**, [S. l.], v. 11, n. 1, p. 1–21, 2021. DOI: 10.1038/s41598-021-81749-9. Disponível em: <https://www.nature.com/articles/s41598-021-81749-9>. Acesso em: 6 jan. 2022.

ZACKS, Michele A.; PAESSLER, Slobodan. ENCEPHALITIC ALPHAVIRUSES. **Veterinary microbiology**, [S. l.], v. 140, n. 3–4, p. 281, 2010. DOI: 10.1016/J.VETMIC.2009.08.023. Disponível em: </pmc/articles/PMC2814892/>. Acesso em: 7 mar. 2022.

ZHANG, Lifang. Multi-epitope vaccines: a promising strategy against tumors and viral infections. **Cellular & Molecular Immunology 2018 15:2**, [S. l.], v. 15, n. 2, p. 182–184, 2017. DOI: 10.1038/cmi.2017.92. Disponível em: <https://www.nature.com/articles/cmi201792>. Acesso em: 6 jan. 2022.

8. ANEXOS

8.1 Sequências das proteínas E2 dos vírus CHIKV e EEEV

>E2_PROTEIN_CHIKV_NP_690589.2

STKDNFNVYKATRPYLAHCPDCGEGHSPVALERIRNEATDGTLKIQVSLQIGIGT
DDSHDWTKLRYMDNHIPADAGRAGLFRVTSAPCTITGTMGHFILARCPKGETLTVGF
TDSRKISHSCTHPFHHDPPVIGREKFHSRPQHGKELPCSTYVQSNAATAEEIEVHMPPD
TPDRILLSQQSGNVKITVNSQTVRYKCNCGGSNEGLITTDKVINNCKVDQCHAAVTN
HKKWQYNSPLVPRNAELGDRKGGKIHIPFPLANVTCMVPKARNPTVTYGKNQVIMLL
YPDHPTLLSYRSMGEEPNYQEEWVTHKKEVVLTVPTGLEVTWGNNEPYKYWPQLS
ANGTAHGHPHEILYYYELYPTMTVVVVSVASFILLSMVGMAVGMCMCARRRCITP
YELTPGATVPFLLSLICCIRTAKA

>E2_PROTEIN_EEEV_NP_632022.1

DLDTHTFTQYKLARPYIADCPNCGHSRCDSPAIIEEVRGDAHAGVIRIQTSAMFGLKTD
GVDLAYMSFMNGKTQKSIKIDNLHVRTSAPCSLVSHHGYYILAQCPPGDTVTVGFHD
GPNRHTCTVAHKVEFRPVGREKYRHPPEHGVELPCNRYTHKRADQGHYVEMHQPG
LVADHSLLSIHSKVKITVPSGAQVKYYCKCPDVREGTTSSDYTTTCTDVKQCRAyli
DNKKWVYNSGRLPRGEGDTFKGKLHVPFVPVKAKCIATLAPEPLVEHKHRTLILHLY
PDHPTLLTTRSLGSDANPTRQWIERPTTVNFTVTGEGLEYTWGNHPPKRVWAQESGE
GNPHGWPHEVVVYYNRYPLTTIIGLCTCVAIIMVSCVTSVWLLCRTRNLCITPYKLA
PNAQVPILLALLCCIKPTRA

8.2 Arquivo FASTA criado a partir do resultado da predição da ferramenta PAP/IMED

>E2_CHIKV_NP_690589.2

n	Start Position	Sequence	End Position
1	7	NVYKATRPYLAHCPDC	22
2	24	EGHSCHSPVAL	34
3	46	LKIQVSLQI	54
4	80	RAGLFVRTSAPCT	92
5	98	GHFILARC	105
6	109	ETLTVGF	115
7	120	KISHSCTHPFHHDPPVIG	137
8	147	HGKELPCSTYVQS	159
9	165	EEIEVHM	171
10	184	QSGNVKITVNSQTVRYKCNC	203
11	216	VINNCKVDQCHAAVT	230
12	237	YNSPLVPR	244
13	253	GKIHIPFPLANVTCMV	268
14	282	QVIMLLYPDHPHTLLSY	297
15	311	VTHKKEVVLTVPTEGLE	327
16	336	YKYWPQL	342
17	349	HGHPHEIILYYYEL	362
18	364	PTMTVVVVSVASFILLSMVGMAVGMCMCARRRCITPYEL	402
19	404	PGATVPFLLSLICCIR	419

>E2_EEEV_NP_632022.1

n	Start Position	Sequence	End Position
1	4	THFTQYKLARPPIADCPNCGHSRCDSPPIAIE	34
2	40	AHAGVIRI	47
3	56	KTDGVDLAY	64
4	77	KIDNLHVRTSAPCSLVSHHGYYILAQCPPGDTVTVGF	113
5	118	NRHTCTVAHKVEFRP	132
6	142	PEHGVELPCNR	152
7	160	QGHYVEM	166
8	168	QPGLVADHSLLSIHSKVKITVPSGAQVKYYCKCPD	203
9	213	YTTTCTDVKQCRAyli	228
10	249	KGKLVFPFVPVKAKCIATLAPEPLVEHKHRTLILHLYPDHPHTLLTT	294
11	331	PPKRVWA	337
12	348	WPHEVVVYYYNR	359
13	361	PLTTIIGLCTCVAIIMVSCVTSVWLLCRTRNLCITPYKLAPNAQVPILLALLCCI	
K	416		

8.3 Arquivo Fasta retornado pela ferramenta POA1 para epítomos da proteína E2 de CHIKV

```
>CHIKV_E2_Bepipred_6_17
FNVYKATRPYLA
>CHIKV_E2_Bepipred_20_30
PDCGEGHSCHS
>CHIKV_E2_Bepipred_36_43
RIRNEATD
>CHIKV_E2_Bepipred_56_80
IGTDDSHDWTKLR YMDNHIPADAGR
>CHIKV_E2_Bepipred_131_167
HDPPVIGREK FHSRPQH GKELPCSTYVQSNAATAEEI
>CHIKV_E2_Bepipred_169_177
VHMPPDTPD
>CHIKV_E2_Bepipred_207_212
NEGLIT
>CHIKV_E2_Bepipred_215_220
KVINNC
>CHIKV_E2_Bepipred_234_252
KWQYNSPLVPRNAELGDRK
>CHIKV_E2_Bepipred_300_314
MGEEPNYQEEWVTHK
>CHIKV_E2_Bepipred_331_350
GNNEPYKYWPQLSANGTAHG
>CHIKV_E2_PAP/IMED_7_22
NVYKATRPYLAHCPDC
>CHIKV_E2_PAP/IMED_24_34
EGHSCHSPVAL
>CHIKV_E2_PAP/IMED_46_54
LKIQVSLQI
>CHIKV_E2_PAP/IMED_80_92
RAGLFVRTSAPCT
>CHIKV_E2_PAP/IMED_98_105
GHFILARC
>CHIKV_E2_PAP/IMED_109_115
ETLTVGF
>CHIKV_E2_PAP/IMED_120_137
KISHSCTHPFHHDPPVIG
>CHIKV_E2_PAP/IMED_147_159
HGKELPCSTYVQS
>CHIKV_E2_PAP/IMED_165_171
EEIEVHM
>CHIKV_E2_PAP/IMED_184_203
QSGNVKITVNSQTVRYKCNC
>CHIKV_E2_PAP/IMED_216_230
VINNCKVDQCHAAVT
...
```

8.4 Alinhamento entre as sequencias das proteínas E2 dos Alphavirus CHIKV e EEEV

```

# Program: needle                                     # Report_file: stdout
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-
#       I20220530-021801-0209-64078438-
#       p2m.asequence
#   -bsequence emboss_needle-
#       I20220530-021801-0209-64078438-
#       p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair

# Aligned_sequences: 2
# 1: E2_CHIKV_NP_690589.2
# 2: E2_EEEV_NP_632022.1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 429
# Identity:      177/429 (41.3%)
# Similarity:   241/429 (56.2%)
# Gaps:         15/429 ( 3.5%)
# Score: 861.0

```

```

E2_CHIKV_NP_6      1 stkdnfnvykatrpylahcpdcgeghs-chspvalerirneatdgtlkiq      49
      .....|..||..|||:|.||:|  ||| |.||:|:|.:.|..|.:||
E2_EEEV_NP_63     1 dldthftqyklarpyiadcpnc--ghsrcdspiaieevrgdahagviriq      48

E2_CHIKV_NP_6     50 vslqigigtddshdwtklrymdnhipadagraglfvrtsapctitgtmgh      99
      .....|:| |..|.....:|:|.....|..|.....|:|.....|:
E2_EEEV_NP_63     49 tsamfglkt-dgvdlaymsfmngktqksikidnlhvrtsapcslvshhgy      97

E2_CHIKV_NP_6     100 filarcpkgetltvgftdsrkishsct--hpfhhdppvigrekfhsrpqh      147
      :||:|:|:|:|:|:|:|..|..|:| |.....| :||:|:|:|:|:|
E2_EEEV_NP_63     98 yilaqcppgdtvtvgfhdgpn-rhtctvahkvefrp--vgrekyrhpph      144

E2_CHIKV_NP_6     148 gkelpcstyvqsnaataaeievhmpdpdtrtllsqsgnvkitvns-qt      196
      |.||||:|.||..|.....:|:|.||..|:|:|..|..|..|..|..|
E2_EEEV_NP_63     145 gvelpcnrythkradqghyvmhqplvadhsllsihsakvkitvpsgaq      194

E2_CHIKV_NP_6     197 vrykcnccgsneglittdkvinnckvdqchaavtnhkkwqynsplvrna      246
      |:|..|.....|:|:|:|.....|..|..|:|:|:|:|:|:|:|:|:|..
E2_EEEV_NP_63     195 vkyyckcpdvregtssdytttctdvkqcraylidnkkwvynsgrlprge      244

E2_CHIKV_NP_6     247 elgdr-kgkihipflanvtcmvpkarnptvtygknqvimllypdhptll      295
      ||. |||:|:|:|.....|:|..|..|:|:|:|:|:|:|:|:|:|:|
E2_EEEV_NP_63     245 --gdtfkgklhvpfvpvkakciatlapeplvehkhrtlilhlypdhptll      292

E2_CHIKV_NP_6     296 syrsmggeepnyqeewthkkevltvpteglevtwgnnepkywpqlsan      345
      :.||:|.:.|.....:|:|.....|..|..|..|..|..|:|:|:|:|:|:|
E2_EEEV_NP_63     293 ttrslgsdanptrqwierpttvnftvtglegleytwgnhppkrvwaqesge      342

E2_CHIKV_NP_6     346 gtahghpheillyyelyptmtvv-vsvasfillsmvgmavgmcmcarr      394
      |..||..|..|:|:|:|..|..|:| :.....|:|:|.....|:|. |.
E2_EEEV_NP_63     343 gnphgwphevvyynryplttiiglctcvaiimvscvtsvllcr-trn      391

E2_CHIKV_NP_6     395 rcitpyeltpgatvpfllsliccirtaka      423
      .||||:|.|.|.||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
E2_EEEV_NP_63     392 lcitpyklapnaqvpillallccikptra      420

#-----
#-----

```

8.5 Alinhamento múltiplo entre as sequencias das proteínas E2 dos Alphavirus CHIKV e EEEV e os epítomos selecionados pela Triagem Otimizada de Antígenos

```

1       10
E2_CHIKV_Epitopes      .....FNVYKATRPYLAHCPDC.....
E2_CHIKV_NP_690589.2  STKDNFNVYKATRPYLAHCPDCEGHSHSCHSPVALERIRNEATDGTLLKIQVSLQIGIGT
E2_EEEV_NP_632022.1  DLDTHFTQYKLARPYIADCPNCGHSR.CDSPIAIEEVRGDAHAGVIRIQTSAMFGLKT
E2_EEEV_Epitopes      .....FTQYKLARPYIADCPNCG.....

20      30      40
E2_CHIKV_Epitopes      .....RAGLFVRTSAPCT.....GHFILARCPKGETLTVGF.
E2_CHIKV_NP_690589.2  DSDSHDWTKLRYMDNHIPADAGRAGLFVRTSAPCTITGTMGHFILARCPKGETLTVGFT
E2_EEEV_NP_632022.1  DGV.DLAYMSFMNGKTQKSIKIDNLHVRTSAPCSLVSHHGYIILAQCPGDTVTVGFH
E2_EEEV_Epitopes      .....LHVRTSAPC.....GYIILAQCPGDTVTVGF.....

E2_CHIKV_Epitopes      .....PQHGKELPCSTYVQSNAAATAEEIEVHMPPD
E2_CHIKV_NP_690589.2  DSRKISHSCTHPFHHDPPVIGREKFHSRPQHGKELPCSTYVQSNAAATAEEIEVHMPPD
E2_EEEV_NP_632022.1  DGNP.RHTCTVAHKVEFRPVGREKYRHPPEHGVELPCNRYTHKRADQGHYVEMHQPGL
E2_EEEV_Epitopes      .....PEHGVELPCNRY.....

50
E2_CHIKV_Epitopes      .....NEGLIT.....
E2_CHIKV_NP_690589.2  TPDRTLLSQSGNVKITVNS.QTVRYKCNCGGSNEGLITTDKVINNCKVDQCHAAVTN
E2_EEEV_NP_632022.1  VADHSLLSIHSAKVKITVPSGAQVKYCKCPDVREGTTSSDYTTTCTDVKQCRAYLID
E2_EEEV_Epitopes      ..DHSLLSIHSAKVKITVPS.....

60      70
E2_CHIKV_Epitopes      .....YNSPLVPR.....KNQVIMLLYP
E2_CHIKV_NP_690589.2  HKKWQYNSPLVPRNAELGDRKGIHIPFPLANVTCMVPKARNPTVTYGKNQVIMLLYP
E2_EEEV_NP_632022.1  NKKWVYNSGRLPR.GEGDTFKGKLVHPFVPVKAACIATLAPEPLVEHKHRTLILHLYP
E2_EEEV_Epitopes      .....ILHLYP.....

80
E2_CHIKV_Epitopes      DHPTLLSYRSMGEPN.....
E2_CHIKV_NP_690589.2  DHPTLLSYRSMGEPNYQEEVWTHKKEVVLTVPTEGLEVTWGNNEPYKYWPQLSANGT
E2_EEEV_NP_632022.1  DHPTLLTTRSLGSDANPTRQWIERPTTVNFTVTGEGLEYTWGNHPKRVWAQESGEGN
E2_EEEV_Epitopes      DHPTLLT.....VTGEGLEYTWGNHP.....

90
E2_CHIKV_Epitopes      .....ITPYELTPG
E2_CHIKV_NP_690589.2  AHGHPHEIILYYEYLYPTMTVVVVSVASFILLSMVGMAGVMCMCARRRCITPYELTPG
E2_EEEV_NP_632022.1  PHGWPHEVVVVYYNRYPLTTIIGLCTCVAIIMVSCVTSVWLLCRTRNLCITPYKLAPN
E2_EEEV_Epitopes      .....LCITPYKLAPN.....

100     110
E2_CHIKV_Epitopes      ATVPFLLSLICCIR....
E2_CHIKV_NP_690589.2  ATVPFLLSLICCIRTAKA
E2_EEEV_NP_632022.1  AQVPILLALLCCIKPTRA
E2_EEEV_Epitopes      AQVPILLALL.....

```

8.6 Epítomos de Linfócitos T auxiliares preditos para mais de um alelo de HLA pela ferramenta MHCII – Binding Predictions

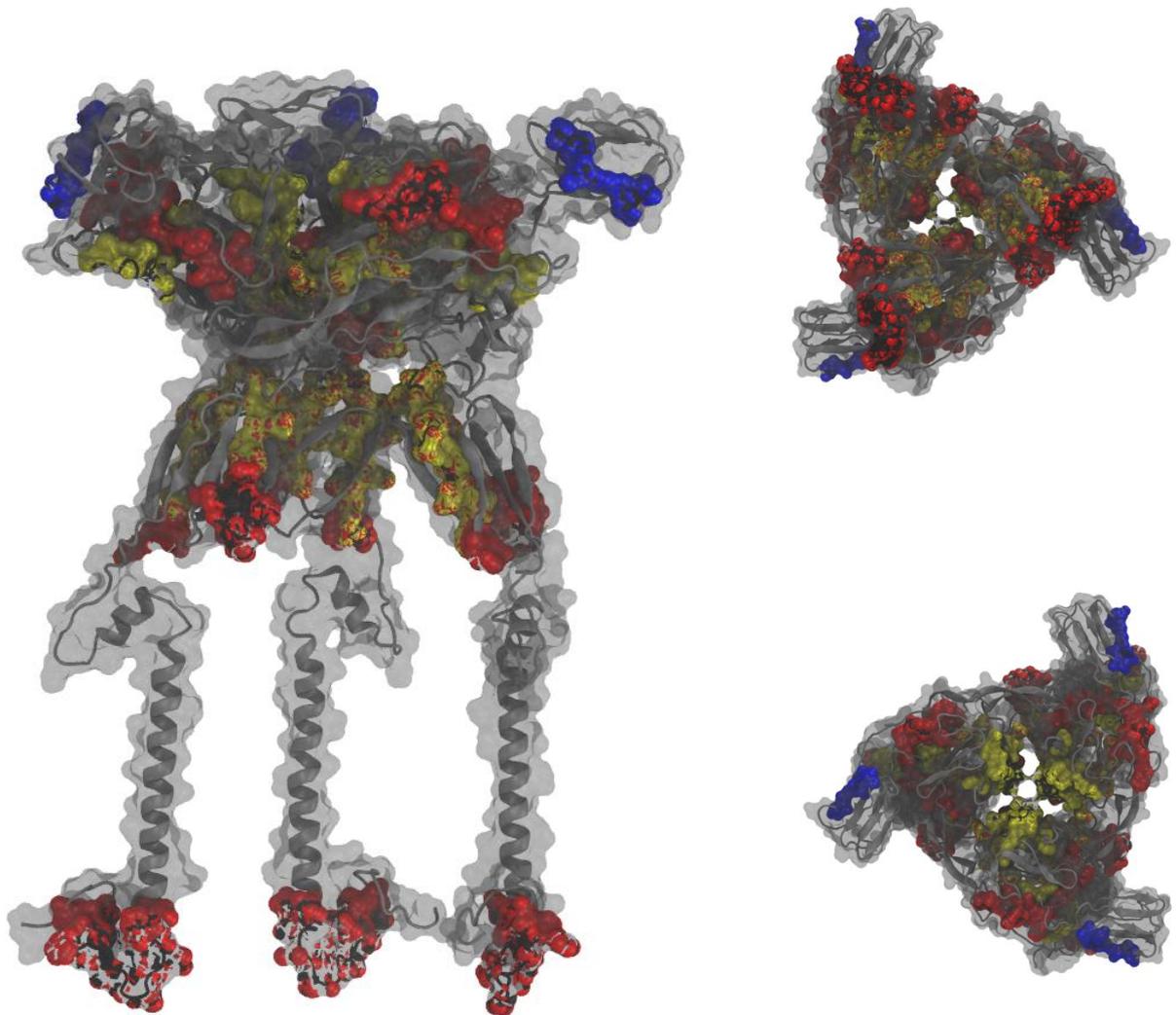
Epítomos de Linfócitos T – auxiliares da proteína E2 conservados (identidade igual ou superior a 60%) entre os Alphavirus CHIKV e EEEV e preditos pela ferramenta MHCII – Binding Predictions com afinidade para mais de um tipo de alelo de HLA.

Espécie	Posição Inicial	Posição Final	Sequência	Alelo de HLA
CHIKV	281	295	NQVIMLLYPDHTLL	HLA-DRB1*15:01
				HLA-DRB4*01:01
	282	296	QVIMLLYPDHTLLS	HLA-DRB1*15:01
				HLA-DRB4*01:01
	6	20	FNVYKATRPYLAHCP	HLA-DRB1*01:01
				HLA-DRB1*07:01
	98	112	GHFILARCPKGETLT	HLA-DRB1*01:01
				HLA-DRB1*11:01
	99	113	HFILARCPKGETLTV	HLA-DRB1*01:01
				HLA-DRB1*11:01
EEEV	6	20	FTQYKLARPVIADCP	HLA-DRB1*01:01
				HLA-DRB1*07:01
				HLA-DRB5*01:01
	174	188	DHSLLSIHSAKVKIT	HLA-DRB1*01:01
				HLA-DRB1*07:01
				HLA-DRB1*13:02
				HLA-DRB5*01:01
	175	189	HSLLSIHSAKVKITV	HLA-DRB1*01:01
				HLA-DRB1*07:01
				HLA-DRB1*13:02
				HLA-DRB5*01:01
	176	190	SLLSIHSAKVKITVP	HLA-DRB1*01:01
				HLA-DRB1*07:01
				HLA-DRB5*01:01
	177	191	LLSIHSAKVKITVPS	HLA-DRB1*01:01
				HLA-DRB1*07:01
	393	407	CITPYKLAPNAQVPI	HLA-DRB1*01:01
				HLA-DRB1*13:02
	394	408	ITPYKLAPNAQVPIL	HLA-DRB1*01:01
				HLA-DRB1*09:01
HLA-DRB1*13:02				
395	409	TPYKLAPNAQVPILL	HLA-DRB1*01:01	
			HLA-DRB1*09:01	
			HLA-DRB1*13:02	

Epítomos de Linfócitos T – auxiliares da proteína E2 conservados (identidade igual ou superior a 60%) entre os Alphavirus CHIKV e EEEV e preditos pela ferramenta MHCII – Binding Predictions com afinidade para mais de um tipo de alelo de HLA (Continuação).

Espécie	Posição Inicial	Posição Final	Sequência	Alelo de HLA
	396	410	PYKLAPNAQVPILLA	HLA-DRB1*01:01
				HLA-DRB1*13:02
	397	411	YKLAPNAQVPILLAL	HLA-DRB1*01:01
				HLA-DRB1*13:02

8.7 Visualização dos epítomos selecionados pela triagem otimizada do POA em estrutura tridimensional da proteína E2 de CHIKV



Visualização 3D da distribuição dos epítomos preditos na proteína E2 de CHIKV, com base no cristal selecionado no banco de dados RCSB PDB (Acesso: 6NK7). Estão caracterizados na estrutura os epítomos lineares preditos para células B pelas ferramentas Bepipred (Azul) e IMED (Amarelo), epítomos preditos para células T-citotóxicas pela ferramenta NetCTL (Verde) e os epítomos de células T-auxiliares, preditos pela ferramenta MHCII Binding Predictions (Vermelho). (A) Vista longitudinal da proteína E2 de EEEV. (B) Vista da proteína E2 rotacionada em 90° para vista superior ao domínio A e B. (C) Vista rotacionada em 90° da proteína E2 sob o domínio C.

8.8 Área exposta ao solvente dos Epítomos de células B selecionados pela Triagem Otimizada de Antígenos

ligand_ID	target_ID	ligand_area	target_area	complex_area	surface_area
CHIKV_207_212_epitope	CHIKV_207_212_protein	999,7434	25.125,6777	25.110,1621	-492,1139
CHIKV_80_92_epitope	CHIKV_80_92_protein	1761,99719	25092,86523	25114,23242	-891,6821899
CHIKV_98_105_epitope	CHIKV_98_105_protein	1353,74023	25565,36523	25111,7207	-450,0478516
CHIKV_109_115_epitope	CHIKV_109_115_protein	1146,21582	25450,93164	25115,35352	-405,3188477
CHIKV_237_244_epitope	CHIKV_237_244_protein	1336,86365	25210,58398	25112,54688	-619,413269
CHIKV_282_297_epitope	CHIKV_282_297_protein	2339,73022	25931,625	25112,5332	-760,3192139
EEEV_323_332_epitope	EEEV_323_332_protein	1.395	24721,80859	24196,05273	-434,3876953
EEEV_142_152_epitope	EEEV_142_152_protein	1679,49426	24256,63281	24195,92969	-809,3955688

